

Universality of the Stochastic Bessel Operator

Brian Rider*, Patrick Waters†

Abstract

We establish universality at the hard edge for general beta ensembles provided that the background potential V is a polynomial such that $x \mapsto V(x^2)$ is uniformly convex and $\beta \geq 1$. The method rests on the corresponding tridiagonal matrix models, showing that their appropriate continuum scaling limit is given by the Stochastic Bessel Operator. As conjectured in [10] and rigorously established in [18], the latter characterizes the hard edge in the case of linear potential and all β (the classical “beta-Laguerre” ensembles).

1 Introduction

We prove a universality result for the limiting distribution of the smallest points for a family of coulomb gas measures. With any $\beta > 0$ and $a > -1$ these measures are prescribed through the joint densities of n points $\{\lambda_1, \dots, \lambda_n\}$ on the positive half-line:

$$c \prod_{i \neq j} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^n w(\lambda_i), \quad w(\lambda) = \lambda^{\frac{\beta}{2}(a+1)-1} e^{-\beta n V(\lambda)}. \quad (1.1)$$

In general, V can be any function that is bounded at zero and of suitable growth at infinity; the constant $c = c(V, \beta, a, n)$ is the corresponding normalizer. The particular choice of the weight w is explained by the fact that when $\beta = 1, 2, 4$, $V(x) = x/2$, and a is an integer, (1.1) is precisely the joint density of eigenvalues for the classical Wishart (or Laguerre) ensembles of random matrix theory. These are ensembles of the form XX^\dagger for an $n \times (n+a)$ matrix X of independent real, complex, or quaternion (at $\beta = 1, 2$, or 4) mean-zero Gaussians, here normalized to have mean-square $(n\beta)^{-1}$.

The scaling limit for the smallest points in this and related contexts is now commonly referred to as the hard-edge limit. In the solvable case of complex Gaussian ensembles ($\beta = 2$ and

*Department of Mathematics, Temple University. e-mail: brian.rider@temple.edu.

†Department of Mathematics, Temple University. e-mail: patrick.waters@temple.edu.

$a = 0, 1, 2, \dots$) a closed form for these distributions was discovered by Tracy and Widom [21], with results for the real and quaternion cases following in [22]. Keeping with the Gaussian-type potential ($V(x) = x^2/2$), but now allowing all possible values of β and a , the densities (1.1) define the extensively studied “beta-Laguerre” ensembles. Based on a corresponding tri-diagonal matrix model of Dumitriu and Edelman [9], Edelman and Sutton [10] conjectured that the limiting beta-Laguerre hard edge should be described by a certain random differential equation which they tagged the Stochastic Bessel Operator. This was subsequently proved in [19].

The Stochastic Bessel Operator ($\text{SBO}_{\beta,a}$) takes the form:

$$\text{SBO}_{\beta,a} = -e^{(a+1)x + \frac{2}{\sqrt{\beta}}b(x)} \frac{d}{dx} e^{-ax - \frac{2}{\sqrt{\beta}}b(x)} \frac{d}{dx}, \quad (1.2)$$

where $x \mapsto b(x)$ is a standard Brownian motion. For the present application, this is viewed as acting on $L^2(\mathbb{R}_+, m(dx))$ for $m(dx) = e^{-(a+1)x - \frac{2}{\sqrt{\beta}}b(x)} dx$ with a Dirichlet boundary condition at the origin. Viewed as a random diffusion generator symmetric with respect to $m(dx)$, one sees that $\text{SBO}_{\beta,a}$ has almost surely discrete spectrum [19].

Here we show that $\text{SBO}_{\beta,a}$ is a universal object, characterizing the hard-edge scaling limit for β -ensembles (1.1) with a certain class of polynomial potentials.

Theorem 1. *Consider the ordered points $0 < \lambda_1 < \lambda_2 < \dots$ drawn from the β -ensemble (1.1) where V is a polynomial such that $x \mapsto V(x^2)$ is uniformly convex and $\beta \geq 1$. Denote by $0 < \Lambda_1 < \Lambda_2 < \dots$ the ordered eigenvalues of $\text{SBO}_{\beta,a}$. Then, there is a constant $c_{V,\beta,a}$ such that: as $n \rightarrow \infty$,*

$$c_{V,\beta,a} n^2(\lambda_1, \lambda_2, \dots) \Rightarrow (\Lambda_1, \Lambda_2, \dots), \quad (1.3)$$

in the sense of finite dimensional distributions.

The restriction to $\beta \geq 1$ is an additional convexity assumption, as we will explain below. Our proof builds on the method in which the $\text{SBO}_{\beta,a}$ limit was originally established in the simpler β -Laguerre setting. We identify a family of tridiagonal matrix models which realize (1.1) as their eigenvalue densities, and then demonstrate that $\text{SBO}_{\beta,a}$ serves as their appropriate continuum operator limit. Along the way we will see that hard-edge universality is a consequence of a certain functional central limit theorem. The whole program is similar to the recent soft-edge universality proof using the characterizing Stochastic Airy Operator [14], as we will also explain in greater detail below.

Hard edge universality has previously been addressed at $\beta = 1, 2, 4$ via the Riemann-Hilbert Problem method: for $\beta = 2$ quite general potentials V are treated in [15], while for $\beta = 1$ and 4 reference [7] considers potentials that are asymptotically monomial. At these values of β the laws (1.1) correspond to eigenvalue densities for nonnegative definite matrices

M drawn according to the law with density proportion to $(\det M)^\gamma e^{-\text{tr} V(M)} dM$ (for choice of $\gamma > -1$). There are also further special values of β (outside of 1, 2, and 4) for which the hard-edge of β -Laguerre can be accessed through multivariate special functions (without appealing to $\text{SBO}_{\beta,a}$), see for example [12]. At the soft edge, besides again $\beta = 1, 2, 4$ results using the Riemann-Hilbert Problem method [6] and the operator approach of [14], there are the results of Bekerman-Figalli-Guionnet [1] and Bourgade-Erdős-Yau [3] which hold for a far more general class of potentials. While these methods most likely extend to the hard edge, the emphasis here (as in [14]) is to demonstrate a simple mechanism for edge universality of random matrices.

Tridiagonals and operator limits

Let $B = B(x, y)$ denote the $n \times n$ lower bi-diagonal matrix

$$B_{i,i} = x_i \text{ for } i = 1, \dots, n, \quad B_{i+1,i} = -y_i \text{ for } i = 1, \dots, n-1, \quad (1.4)$$

with the convention that all x_i and y_i are positive. Build the random $B = B(X, Y)$ with the variables $(X_1, \dots, X_n, Y_1, \dots, Y_{n-1})$ drawn according to the density,

$$P(x_1, \dots, x_n, y_1, \dots, y_{n-1}) = c \exp [-n\beta \text{tr} V(BB^T)] \prod_{k=1}^n x_k^{\beta(k+a)-1} \prod_{k=1}^{n-1} y_k^{\beta k-1}, \quad (1.5)$$

on $(\mathbb{R}_+)^{2n-1}$. Then, the key fact is that the random tridiagonal $B(X, Y)B(X, Y)^T$ has joint eigenvalue density given by (1.1). This is the general potential analogue of the Edelman-Dumitriu result [9]. When V is linear (1.5) reduces to their representation of β -Laguerre: all X_i and Y_i are independent with $X_i \sim \frac{1}{\sqrt{n\beta}} \chi_{\beta(i+a)}$ and $Y_i \sim \frac{1}{\sqrt{n\beta}} \chi_{\beta i}$ (χ_r denoting a chi variable of parameter $r > 0$). The proof is much the same as that in [9], and for completeness is included in the appendix.

Next we recall (from [19]) that $\text{SBO}_{\beta,a}$ is best understood through its inverse, which has a similar decomposition to the the matrix model $(BB^T)^{-1}$. Mapped to act on $[0, 1]$ rather than the half line, this inverse takes the form $K^T K$ in which K is the integral operator with kernel¹

$$K(s, t) = \frac{1}{\sqrt{t}} \left(\frac{s}{t} \right)^{a/2} \exp \left[\int_s^t \frac{db_u}{\sqrt{\beta u}} \right] 1_{s < t} \quad (1.6)$$

on $L^2[0, 1]$. The strategy that emerges is to show that, after an embedding into $L^2[0, 1]$, $[nB]^{-1}$ converges to K in a suitably strong sense.

¹Throughout we use the same notation for any integral operator and its corresponding kernel.

Now to be completely concrete we specify $V(x) = \sum_{m=1}^d g_m x^m$, and introduce $t \mapsto \phi(t)$ as the unique positive solution of

$$t = \sum_{m=1}^d m \binom{2m}{m} g_m \phi(t)^{2m}, \quad \text{for } t \in [0, 1]. \quad (1.7)$$

In terms of ϕ , we also define

$$\theta(t) = \kappa \left(\int_0^t \frac{du}{\phi(u)} \right)^2, \quad \text{with } \kappa = \kappa_{V,\beta,a} \text{ chosen so that } \theta(1) = 1. \quad (1.8)$$

That ϕ has the claimed properties and is such that θ is finite requires some verification. Granted this however our main technical result is the following.

Theorem 2. *Let $x \mapsto V(x^2)$ be a uniformly convex polynomial and take $\beta \geq 1$. Denote by K_n the canonical embedding of the random matrices $[nB(X, Y)]^{-1}$ as operators from $L^2[0, 1]$ to itself. Then, for any sequence $n \rightarrow \infty$ there is a subsequence $n' \rightarrow \infty$ and a probability space on which $K_{n'}$ converges to the integral operator K with kernel*

$$K(s, t) = \frac{1}{\sqrt{\phi(s)\phi(t)}} \left(\frac{\theta(s)}{\theta(t)} \right)^{\frac{a}{2} + \frac{1}{4}} \exp \left[\frac{1}{\sqrt{\beta}} \int_{\theta(s)}^{\theta(t)} \frac{db_z}{\sqrt{z}} \right] 1_{s < t}, \quad (1.9)$$

almost surely in Hilbert-Schmidt norm.

One observes that when $V(x) = x/2$, the definitions (1.7) and (1.8) yield $\phi(t) = \sqrt{t}$ and $\theta(t) = t$, and (1.9) reduces to the advertised kernel K in (1.6). In general we have that

$$K(\theta(s), \theta(t)) \sqrt{\theta'(s)\theta'(t)} = 2\kappa K(s, t), \quad (1.10)$$

with κ as in (1.8). In other words, the eigenvalues of $K^T K$ and $K^T K$, defined with the same Brownian motion, agree up to an overall multiple of $4\kappa^2$, identifying the (V, β, a) -dependent scaling constant $c_{V,\beta,a} = 4\kappa^2$ in (1.3). Here we are using, as is implicit the statement of Theorem 2, that K and K are almost-surely Hilbert-Schmidt. To conclude Theorem 1 is more or less immediate. In the subsequential coupling of Theorem 2 one has $K_n^T K_n \rightarrow K K^T$ in trace norm. Hence, the finite parts of the spectrum of $(nB B^T)^{-1}$ converge to those of $K^T K$ in this manner and so also in distribution. In particular we have as well the convergence in distribution of (any fixed number) of the eigenfunctions as elements of $L^2[0, 1]$.

Remark 1. Rather than embedding $[nB]^{-1}$ according to the “flat” basis $e_k = \sqrt{n} 1_{[(k-1)/n, k/n]}$ and performing the change of variables (1.10) after the fact, we could work instead with the suitably weighted $1_{[\theta(k-1/n), \theta(k/n)]}$ basis functions to define the embedding. Then, after scaling by $4\kappa^2$, the corresponding discrete kernels will converge to K itself.

Remark 2. The introduced function ϕ turns out to provide a first order approximation for the appropriate energy function (Hamiltonian) associated with P . It can also be described through a “time-dependent” version of the equilibrium measure for the eigenvalue law (1.1). In particular, let $V_t = t^{-1}V$ for $t \in (0, 1]$ and consider

$$\mu_t = \operatorname{argmin}_{\mu \in M} \int_0^\infty V_t(x) \mu(dx) - \int_0^\infty \int_0^\infty \log |x - y| \mu(dx) \mu(dy),$$

where M is the space of probability measures of the half-line. It is the case that μ_t has support $[0, \phi(t)]$.

Overview of the proof

Using the explicit inversion formula for bidiagonal matrices, the basic object of study is now understood to be the random kernel operator

$$K_n(s, t) = \frac{1}{X_j} \prod_{k=i}^{j-1} \frac{Y_k}{X_k} 1_{\Gamma_{ij}}(s, t). \quad (1.11)$$

Here Γ_{ij} is the set on which $s \in [\frac{i-1}{n}, \frac{i}{n})$, $t \in [\frac{j-1}{n}, \frac{j}{n})$, and $s < t$, when $i = j$ the product in (1.11) is understood to equal one. Given this expression, that $\operatorname{spec}([nBB^T]^{-1}) = \operatorname{spec}(K_n^T K_n)$ can be checked by hand.

The measure P under which K_n is drawn has the form $\frac{1}{Z} e^{-n\beta H} dx dy$ with Hamiltonian

$$H(x, y) = \operatorname{tr} V(BB^T) - \sum_{k=1}^n \left(\frac{k}{n} + \frac{a}{n} - \frac{1}{n\beta} \right) \log x_k - \sum_{k=1}^{n-1} \left(\frac{k}{n} - \frac{1}{n\beta} \right) \log y_k. \quad (1.12)$$

Our assumptions imply that P is uniformly log-concave, that is $(\nabla^2 H)(x, y) \geq cI$ for some $c > 0$ and all $(x, y) \in \mathbb{R}_+^{2n-1}$. In particular, with $\beta \geq 1$ each of the log terms in (1.12) has nonnegative second derivative. One then concludes by noting that,

$$\operatorname{tr} V(BB^T) = \frac{1}{2} \operatorname{tr} V(A^2), \quad \text{for } A = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}.$$

and applying C. Davis’ theorem [5]: a (uniformly) convex function of a Hermitian matrix is (uniformly) convex as a function of its entries. Since (1.11) is a simple functional of the process $k \mapsto (X_k, Y_k)$, one is left to quantify the anticipated Gaussian fluctuations of (X_k, Y_k) about the minimizer (x_k^o, y_k^o) of the Hamiltonian H .

In Section 2 we develop a fine (out to $o(n^{-1})$) approximation of the minimizer which allows us to establish the correct centering:

$$\lim_{n \rightarrow \infty} \sum_{k=\lfloor ns \rfloor}^{\lfloor nt \rfloor} \log \frac{x_k^o}{y_k^o} = - \left(\frac{a}{2} + \frac{1}{4} \right) \log \frac{\theta(t)}{\theta(s)} + \frac{1}{2} \log \frac{\phi(t)}{\phi(s)}, \quad (1.13)$$

for all fixed s, t with $0 < s < t < 1$. Granted this, the limiting kernel (1.9) is identified by showing that

$$X_{[nt]} \Rightarrow \phi(t), \quad \sum_{k=[nt]}^n \log \frac{X_k/x_k^o}{Y_k/y_k^o} \Rightarrow \frac{1}{\sqrt{\beta}} \int_{\theta(t)}^1 \frac{db_u}{\sqrt{u}} \quad (1.14)$$

in the Skorohod topology on $(0, 1]$. Here the polynomial assumption on V is important as it gives P a Markov field property: for example, (X_i, Y_i) and (X_j, Y_j) with $|i - j| > d$ are conditionally independent given any intervening block of variables of length d . The implied decorrelation is quantified in a deterministic way, by showing a decay of dependence of the minimizers of conditional versions of the Hamiltonian H with respect to boundary conditions. These estimates also appear in Section 2. Section 3 builds up further properties of the measure P , in particular demonstrating Gaussian concentration about the minimizer in terms of tail properties as well as Gaussian approximation of the expectation of various test functions. Given all this, the proof of (1.14) appears in Section 4.

Together (1.13) and (1.14) provide point-wise convergence (in law) of K_n to K . To prove that $\int_0^1 \int_0^1 |K_n - K|^2 \rightarrow 0$ (over subsequences) as claimed in Theorem 2 requires a certain domination of K_n by a tight family of $L^2([0, 1]^2)$ kernels. This is carried out in Section 5.

Comparison with the soft edge

Before getting on with it, we have a few comments on the technical differences between the present result and the program carried out in [14] for the soft edge. While the Stochastic Airy Operator is a more delicate object than our integral operator K , being a differential operator and so “local”, to understand the underlying operator convergence of the tridiagonal models one only requires fine information on the first $O(n^{1/3})$ entries of those matrices. For the hard edge we have in a sense to resolve $O(n)$ of n variables. This requires a much more elaborate estimate on the minimizers of H , as well as a better handle on the decorrelation between separate stretches of variables under P . While our functional central limit theorem in Section 4 follows a standard blocking strategy, the corresponding calculation in [17] is really a “one-block” estimate. This also in part explains why the allied method in [14] works for all $\beta > 0$. For us the issue is near “the singularity”, or for X and Y of small index where the measure P becomes less coercive. Or, said another way: where, when $\beta < 1$, the Hamiltonian fails to be convex. While the same issue appears in the soft edge, the troublesome indices are beyond the $O(n^{1/3})$ cutoff and one can get by with rather rough estimates on that part of the field. Again, for the present calculation we simply need more precise control of these entries (as evidenced in particular in the operator norm estimates of Section 5). Assuming $\beta \geq 1$ covers the classical cases, and makes an already technical paper a little less so.

2 Minimizers

As indicated above, the function ϕ introduced in (1.7) serves as a first order approximation to the minimizer (x^o, y^o) of the Hamiltonian H . The idea is to focus about a fixed index k at a continuum position $t = k/n$, ignoring the additional $(a/n - 1/n\beta)$ and $(-1/n\beta)$ multipliers of the $\log x_k$ and $\log y_k$ terms in (1.12). Then from the potential we keep only those terms in $\text{tr } V(MM^T)$ which involve (x_k, y_k) . Assuming that the minimizer is locally constant (*i.e.*, we posit all x_ℓ and y_ℓ for $|\ell - k| \leq d$ equal some x and y), we arrive at the following “coarse Hamiltonian” (at fixed $t \in [0, 1]$):

$$H_t(x, y) = \sum_{m=1}^d g_m \sum_{\ell=0}^m \binom{m}{\ell}^2 x^{2\ell} y^{2m-2\ell} - t \log x - t \log y. \quad (2.1)$$

Plainly H_t is symmetric in x and y , its common “coarse minimizer” $x(t) = y(t)$ defining $\phi(t)$. Note that the relation (1.7) is simply the equation for a critical point of H_t .

A similar approximation was employed in [14] where the analogous objects are referred to as the local Hamiltonian and corresponding local minimizer. Here though we require a sharper approximation. In particular, to pin down the limiting mean of the log potential, recall (1.13), one must refine (x^o, y^o) to $o(n^{-1})$ errors.

Definition: In terms of ϕ define the functions $t \mapsto x^{(1)}(t)$ and $t \mapsto y^{(1)}(t)$ via:

$$\begin{aligned} x^{(1)}(t) - y^{(1)}(t) &= \left(a + \frac{1}{2}\right) \left(\int_0^t \frac{du}{\phi(u)}\right)^{-1} - \frac{\phi'(t)}{2}, \\ x^{(1)}(t) + y^{(1)}(t) &= \left(a - \frac{2}{\beta}\right) \phi'(t). \end{aligned} \quad (2.2)$$

Then, for $i \in [1, n]$, set

$$x_i^\sharp = \phi(i/n) + \frac{x^{(1)}(i/n)}{n}, \quad y_i^\sharp = \phi(i/n) + \frac{y^{(1)}(i/n)}{n}. \quad (2.3)$$

We will refer to (x^\sharp, y^\sharp) as the “fine minimizer”.

Proposition 8 proved below in this section shows that for bulk indices $|x_i^o - x_i^\sharp|$ and $|y_i^o - y_i^\sharp|$ are in fact $O(n^{-2})$, from which the desired appraisal (1.13) follows.

The identification of (x^\sharp, y^\sharp) relies on uniform convexity in an essential way. Indeed, a characterization equivalent to $\text{Hess } H(x, y) \geq cI$ for all $(x, y) \in \mathbb{R}_+^{2n-1}$ is that

$$c\|(x, y) - (x', y')\|_2^2 \leq \left\langle \nabla H(x, y) - \nabla H(x', y'), (x, y) - (x', y') \right\rangle, \quad (2.4)$$

for all (x, y) and (x', y') . Putting $(x', y') = (x^o, y^o)$, the undetermined true minimizer, and applying the Cauchy-Schwartz inequality (2.4) implies that

$$\|(x, y) - (x^o, y^o)\|_2 \leq \frac{1}{c} \|\nabla H(x, y)\|_2. \quad (2.5)$$

The point is that the fine minimizer (x^{\sharp}, y^{\sharp}) has been engineered so that $\nabla H(x^{\sharp}, y^{\sharp})$ vanishes to sufficiently high order.

2.1 Identifying the fine minimizer

The goal of this subsection is to establish the following.

Lemma 3. *For any $i \in [1, n - d)$ it holds that*

$$\left| \frac{\partial}{\partial x_i} H(x^{\sharp}, y^{\sharp}) \right| + \left| \frac{\partial}{\partial y_i} H(x^{\sharp}, y^{\sharp}) \right| \leq c \frac{1}{\sqrt{ni^3}} \quad (2.6)$$

with a constant c depending on V, β and a . For $i \in [n - d, n]$ the right hand side of (2.6) can be replaced by $O(1)$.

Before this however we go back and verify that the coarse minimizer has the various properties claimed above, and also provide an estimate on its shape near the singularity which will be needed for the proof of Lemma 3.

Lemma 4. *The coarse minimizer $\phi(t)$ is unique, positive and increasing, and satisfies*

$$c^{-1} \leq \phi(t)t^{-1/2}, \quad \phi'(t)t^{1/2}, \quad \phi''(t)t^{3/2} \leq c \quad (2.7)$$

for a constant c and all small enough $t > 0$.

Proof. Uniqueness can be seen from the following alternative description of H_t . Let C be the circulant version of our bidiagonal matrix B in variables x_1, \dots, x_m and y_1, \dots, y_m (with $m > d$) and consider minimizing

$$(x, y) \mapsto \text{tr } V(CC^T) - t \sum_{k=1}^m (\log x_k + \log y_k). \quad (2.8)$$

By another application of C. Davis' theorem, this is a convex function. It is also invariant under rotations of the indices, and so its minimizer satisfies $x_k \equiv x$ and $y_k \equiv y$ for some x and y and all k . But making this substitution one finds that the right hand side of (2.8) equals $mH_t(x, y)$.

Showing that $t \mapsto \phi(t)$ is positive and increasing comes down to showing that the right hand side of (1.7) is increasing as a function of ϕ . Rewrite that expression as in

$$\sum_{m=1}^d m \binom{2m}{m} g_m \phi^{2m} = \phi \int_0^2 2u \phi V'(\phi^2 u^2) \frac{4u du}{\sqrt{4 - u^2}}. \quad (2.9)$$

Since $x \mapsto V(x^2)$ is uniformly convex, $x \mapsto 2xV'(x^2)$ is increasing. The integral on the right hand side of (2.9) is therefore a weighted average of such increasing functions, which yields the claim.

As for (2.7), again by uniform convexity we have that $g_1 > 0$, and so for small t the relation (1.7) takes the form $t = 2g_1\phi^2(1 + o(1))$. This shows that $\phi(t)$ is bounded above and below by a multiple of \sqrt{t} for small t . The estimates on ϕ' and ϕ'' follow suit by considering successive derivatives of (2.7). \square

Proof of Lemma 3. The starting point is a lattice path representation for the diagonal entries of powers of BB^T : for $i \in (d, n - d)$,

$$[(BB^T)^m]_{ii} = \sum_{p \in P_m} \prod_{j=1}^m \left(\begin{cases} x_{i+p(2j-1)} & \text{if } p(2j) = p(2j-1) \\ -y_{i+p(2j-1)-1} & \text{if } p(2j) = p(2j-1) - 1 \end{cases} \right) \times \left(\begin{cases} x_{i+p(2j)} & \text{if } p(2j+1) = p(2j) \\ -y_{i+p(2j)} & \text{if } p(2j+1) = p(2j) + 1 \end{cases} \right). \quad (2.10)$$

Here P_m denotes the collection of random walk paths of length $2m$ beginning and ending at height 0 and constrained as follows. At odd-timed steps j (corresponding to selecting an entry from B), the path can either take a step of type \rightarrow and remain at $p(j)$ or take a step of type \searrow and $p(j+1) = p(j) - 1$. At even-timed steps j (corresponding to selecting an entry from B^T) the step can either be again of type \rightarrow or of type \nearrow , in which case $p(j+1) = p(j) + 1$.

Note that for $i \in [1, d] \cup [n - d, n]$ certain paths will be truncated, resulting in a more cumbersome expression.

From (2.10) it is easy to see that: with again $i \in (d, n - d)$,

$$\frac{\partial}{\partial x_i} \text{tr} V(BB^T) = \sum_{m=1}^d g_m \sum_{r=1}^{2m} \frac{r}{x_i} \sum_{p \in P_{m,r}} \prod_{j=1}^m \left[\left(\begin{cases} x_{i+p(2j-1)} & \text{if step } 2j-1 \text{ is } \rightarrow \\ y_{i+p(2j-1)-1} & \text{if step } 2j-1 \text{ is } \searrow \end{cases} \right) \times \left(\begin{cases} x_{i+p(2j)} & \text{if step } 2j \text{ is } \rightarrow \\ y_{i+p(2j)} & \text{if step } 2j \text{ is } \nearrow \end{cases} \right) \right], \quad (2.11)$$

where now $P_{m,r}$ is the subset of P_m comprised of those paths that posses exactly r steps of type \rightarrow at height zero.

The next step is to substitute the values of the fine minimizer into (2.11). These are used in the form: for $|k| \leq d$ and $i + k \in [1, n]$,

$$x^{\sharp}(i+k) = \phi(i/n) + \frac{1}{n} \left(k\phi'(i/n) + x^{(1)}(i/n) \right) + O\left(\frac{1}{\sqrt{ni^3}}\right), \quad (2.12)$$

with a like expression for $y^{\sharp}(i+k)$. To see (2.12), $\phi(t+k/n)$ and $x^{(1)}(t+k/n)$ are expanded out to second and first order, respectively. That both $(d^2/dt^2)\phi(t)$ and $(d/dt)x^{(1)}(t)$ are

$O(t^{-3/2})$ follows from Lemma 4. The result of the substitution is:

$$\begin{aligned} & \frac{\partial}{\partial x_i} \text{tr} V(BB^T)(x^{\frac{1}{2}}, y^{\frac{1}{2}}) \\ &= \sum_{m=1}^d g_m \left[A_m \phi^{2m-1} + \frac{\phi^{2m-2}}{n} (B_m x^{(1)} + C_m y^{(1)} + D_m \phi') \right] + O\left(\frac{1}{\sqrt{ni^3}}\right), \end{aligned} \quad (2.13)$$

where the functions $\phi, \phi', x^{(1)}$ and $y^{(1)}$ are all evaluated at i/n , and

$$\begin{aligned} A_m &= m \binom{2m}{m}, & B_m &= \frac{2m^2 - 2m + 1}{2m - 1} A_m, \\ C_m &= \frac{2m^2 - 2m}{2m - 1} A_m, & D_m &= -\frac{m^2 - m}{2m - 1} A_m. \end{aligned} \quad (2.14)$$

Putting off the derivations behind (2.14) we can complete the proof.

For the x_i -derivative of the logarithmic term in the Hamiltonian we have that,

$$\left(\frac{i + a - \beta^{-1}}{n} \right) \frac{1}{x^{\frac{1}{2}}(i)} = \frac{i + a - \beta^{-1}}{n\phi(i/n)} - \frac{ix^{(1)}(i/n)}{n^2\phi^2(i/n)} + O\left(\frac{1}{\sqrt{ni^3}}\right). \quad (2.15)$$

And if we instead consider $\frac{\partial H}{\partial y_i}$ we arrive at formulas similar to (2.13) and (2.15) where: in the analog of the former, B_m and C_m change roles and D_m changes sign, while in the latter $y^{(1)}$ replaces $x^{(1)}$ on the right hand side and a is set to zero. The claim then is that the formulas for $x^{(1)}$ and $y^{(1)}$ are equivalent to:

$$\begin{aligned} \begin{pmatrix} a - \beta^{-1} - \frac{t}{\phi} x^{(1)} \\ -\beta^{-1} - \frac{t}{\phi} y^{(1)} \end{pmatrix} &= \sum_{m=1}^d \frac{g_m m}{2m - 1} \binom{2m}{m} \phi^{2m-1} \\ &\quad \times \begin{pmatrix} 2m^2 - 2m + 1 & 2m^2 - 2m & -m^2 + m \\ 2m^2 - 2m & 2m^2 - 2m + 1 & m^2 - m \end{pmatrix} \begin{pmatrix} x^{(1)} \\ y^{(1)} \\ \phi' \end{pmatrix}, \end{aligned} \quad (2.16)$$

in which $t = i/n$, the A_m terms having dropped out by the definition (1.7) of ϕ . Now adding the rows of (2.16) we find that the expression for $x^{(1)}(t) + y^{(1)}(t)$ from (2.2) would follow from the identity

$$\sum_{m=1}^d \frac{g_m m (4m^2 - 4m + 1)}{2m - 1} \binom{2m}{m} \phi(t)^{2m-1} = \frac{1}{\phi'(t)} - \frac{t}{\phi(t)},$$

but this is another consequence of (1.7). The expression for $x^{(1)}(t) - y^{(1)}(t)$ is checked in a similar way.

We remark that while this exact vanishing cannot hold for indices $i \in [1, d]$ or $[n-d, n]$, for $i = O(1)$ Lemma 4 implies that the right hand sides of both (2.13) and (2.15) are $O(1/\sqrt{n})$ allowing the estimate (2.6) to be extended to the lower range.

We now go back and derive the coefficients (2.14). In counting weighted $P_{m,r}$ paths it is convenient to introduce the following bijection. Denote by $\tilde{P}_{m,r}$ those paths in $P_{m,r}$ for which the first step is of type \rightarrow , and define

$$(p, j) \in P_{m,r} \times \{1, \dots, r\} \mapsto (\tilde{p}, q) \in \tilde{P}_{m,r} \times \{0, \dots, 2m-1\}, \quad (2.17)$$

where \tilde{p} is obtained by shifting p to the left so the the j^{th} height-zero step of type \rightarrow of p becomes the first step of \tilde{p} . The number q tracks how far p is shifted to produce \tilde{p} .

For A_m we select the same factor (the $\phi(i/n)$ from (2.12)) in each factor in the inner product of (2.13), and so

$$A_m = \sum_{r=1}^{2m} r |P_{m,r}| = 2m |\tilde{P}_m| = 2m \sum_{\ell=0}^m \binom{m-1}{\ell} \binom{m}{\ell}. \quad (2.18)$$

The second equality uses (2.17) with \tilde{P}_m denoting the union over all $\tilde{P}_{m,r}$, that is, those paths just constrained to start with a \rightarrow step. Then we sum over choices of positions for the ℓ steps of type \searrow among the remaining $m-1$ odd-timed steps, balanced by a choice of ℓ (of m possible) steps of type \nearrow at the even-timed steps. The last expression in (2.18) can be written $2 \sum_{\ell=0}^m \ell \binom{2m}{\ell}^2$ after a change of variable which in turn equals $m \binom{2m}{m}$.

In computing B_m , one of the previous ϕ factors is now a $x^{(1)}$. These can only appear at \rightarrow steps, and so in the sum over $P_{m,r}$ paths one has the weight $(\#\{\rightarrow \text{ steps}\} - 1)$ to account for the possible choices of position of the $x^{(1)}$ factor. (The -1 shift is due the fact we have differentiated in the x_i variable.) Hence, with an obvious shorthand and by the same reasoning behind (2.18):

$$B_m = \sum_{r=1}^{2m} r \sum_{p \in P_{m,r}} (\#\rightarrow - 1) = 2m \sum_{p \in \tilde{P}_m} (\#\rightarrow - 1) = 2m \sum_{\ell=0}^m (2m - 2\ell - 1) \binom{m-1}{\ell} \binom{m}{\ell}.$$

Similar to before we can rewrite the above as $\sum_{\ell=0}^m (2\ell)(2\ell-1) \binom{2m}{\ell}^2$, from which the expression in (2.14) follows from the derivation of A_m along with known expressions for $\sum_{\ell=0}^m \ell^2 \binom{2m}{\ell}^2$. The calculation for C_m is basically the same, with $\#\nearrow \cup \searrow$ in place of $\#\rightarrow - 1$.

Finally turning to D_m , note that any appearance of ϕ' is weighted by the relative height of the path, and we have that

$$\begin{aligned} D_m &= \sum_{r \geq 0} \sum_{p \in P_{m,r}} r \sum_{j=1}^m \left(p(2j-1) + \begin{cases} p(2j) & \text{if step } 2j \text{ is } \rightarrow \\ p(2j) - 1 & \text{if step } 2j \text{ is } \nearrow \end{cases} \right) \\ &= \sum_{r \geq 0} \sum_{p \in P_{m,r}} r \sum_{j=1}^m 2p(2j-1) = 4m \sum_{p \in \tilde{P}_m} \sum_{j=1}^m p(2j-1), \end{aligned} \quad (2.19)$$

where we have used that the corresponding weight is always the smaller of the heights across any step. To evaluate the last sum in (2.19) we use a method which we learned from [11] (see in particular Prop. 4.2):

$$\begin{aligned}
\sum_{p \in \tilde{P}_m} \sum_{j=1}^m p(2j-1) &= \sum_{\substack{i_1, k_1, m_1, i_2, k_2, m_2 \geq 0 \\ m_1 + m_2 = m \\ i_1 - k_1 + i_2 - k_2 = 0}} (i_1 - k_1) \binom{m_1 - 1}{i_1} \binom{m_1 - 1}{k_1} \binom{m_2 + 1}{i_2} \binom{m_2}{k_2} \\
&= [w^0][z^m] \left(\sum_{i_1, k_1, m_1, i_2, k_2, m_2 \geq 0} u \partial_u u^{i_1 - k_1} v^{i_2 - k_2} z^{m_1 + m_2} \right. \\
&\quad \left. \times \binom{m_1 - 1}{i_1} \binom{m_1 - 1}{k_1} \binom{m_2 + 1}{i_2} \binom{m_2}{k_2} \right) \Big|_{u, v \rightarrow w} \\
&= [w^0][z^m] \left(u \partial_u \frac{z}{1 - z(u + 2 + u^{-1})} \frac{1 + v}{1 - z(v + 2 + v^{-1})} \right) \Big|_{u, v \rightarrow w}.
\end{aligned}$$

In line one, i_1 and k_1 count the running number of type \nearrow and \searrow steps, respectively. In line two, we have used the notation $[x^p]f(x)$ for the p^{th} coefficient of the Taylor expansion of the (analytic) function $x \mapsto f(x)$. The remaining evaluations are straightforward. \square

2.2 Minimizers and boundary conditions

Here we consider the Hamiltonian H subject to certain boundary conditions. To be more precise, start by fixing an interval $I = [i_0, i_1] \subset [1, n]$ and denote by ∂I (the boundary of I) the at most d indices to the left/right of I . That is, $\partial I = ([i_0 - d, i_0 - 1] \cup [i_1 + 1, i_1 + d]) \cap [1, n]$. View H as a function of $(x, y) \in I$ with those coordinates whose indices lie in ∂I prescribed to equal some values q . By the assumptions on V , those $(x, y) \in I \cup \partial I$ decouple from the $(x, y) \in I$ given q . This restricted function is referred to as the “conditional Hamiltonian” H_q with boundary conditions q .

The goal is to quantify at what rate the minimizers of H_q become independent of q as one moves away from the boundary.

Proposition 5. *For an interval $I \subset [1, n]$ consider the conditional Hamiltonian H_q , i.e., H restricted to I with the coordinates in ∂I set equal to some values q . Assume $\|q\|_\infty \leq c'$. Then, with (x^q, y^q) the minimizer of H_q , it holds that*

$$|x_i^q - x_i^o| + |y_i^q - y_i^o| \leq c \|q - (x^o, y^o)\|_{\infty, \partial I} e^{-\text{dist}(i, \partial I)/c}, \quad (2.20)$$

for any $i \in I$. Here $c = c(V, \beta, a, c')$. If $I = [i_0, i_1] \subset [1, n - d]$ there is also the bound,

$$|x_i^q - x_i^{\sharp}| + |y_i^q - y_i^{\sharp}| \leq c \max \left(\frac{1}{\sqrt{ni_0^3}}, \|q - (x^{\sharp}, y^{\sharp})\|_{\infty, \partial I} e^{-\text{dist}(i, \partial I)/c} \right), \quad (2.21)$$

for any $i \in I$.

Proposition 5 is a deterministic version of decorrelation, and will play an important role in the blocking estimates behind the functional central limit theorem in Section 4. More presently it is used to turn the calculation of Lemma 3 into a fairly optimal estimate on the distance between the true and fine minimizers (Proposition 8 in the next subsection).

The proof of Proposition 5 is based on the following two lemmas. The first, Lemma 6, is a direct consequence of the uniform convexity criteria (2.4) - (2.5). This is then bootstrapped to yield the proposition with the help of Lemma 7, which is a kind of discrete Gronwall inequality. The proof of the latter is a simple inductive argument which is not reproduced here.

The program is quite similar to that in Sections 6-7 of [14]. However, the use of (2.4) - (2.5) streamlines things considerably, bypassing for example the a priori lower bounds on minimizers required in [14].

Lemma 6. *For any conditional minimizer (x^q, y^q) of an H_q defined on some $I \subset [1, n]$,*

$$\|(x^q, y^q) - (x^o, y^o)\|_{2,I}^2 \leq \rho(q) \|q - (x^o, y^o)\|_{2,\partial I}^2 \quad (2.22)$$

where $\rho(q)$ is polynomial of degree $2d$ in the boundary variables q (with bounded coefficients depending only on V, β , and a). And if $I \subset [1, n-d]$ it also holds that

$$\|(x^q, y^q) - (x^{\sharp}, y^{\sharp})\|_{2,I}^2 \leq \sum_{i \in I} \frac{c}{ni^3} + \rho(q) \|q - (x^{\sharp}, y^{\sharp})\|_{2,\partial I}^2 \quad (2.23)$$

with another polynomial ρ of degree $2d$ and $c = c(V, \beta, a)$.

Lemma 7. *Let a_i and b_i be nonnegative sequences satisfying $\sum_{i=0}^k a_i \leq ca_{k+1} + \sum_{i=0}^k b_i$ for a constant c and all $k \leq m$. Then it holds that*

$$a_0 \leq c \left(\frac{c}{c+1} \right)^k a_{k+1} + \sum_{i=0}^k \left(\frac{c}{c+1} \right)^i b_i,$$

again for all $k \leq m$.

Proof of Lemma 6. We first observe that (x^o, y^o) is bounded in sup-norm, independent of the dimension. By (2.5): with a different constant c ,

$$\|(x^{\sharp}, y^{\sharp}) - (x^o, y^o)\|_{\infty}^2 \leq \|(x^{\sharp}, y^{\sharp}) - (x^o, y^o)\|_2^2 \leq c \|\nabla H(x^{\sharp}, y^{\sharp})\|_2^2,$$

and this is $O(1)$ by Lemma 3. The explicit formulas (2.2) and (2.3) then show that $\|(x^{\sharp}, y^{\sharp})\|_{\infty}$ is bounded independently of n which yields the claim.

Similarly,

$$\|(x^o, y^o) - (x^q, y^q)\|_{I,\infty}^2 \leq c \sum_{i \in I: \text{dist}(i, \partial I) \leq d} \left(|\partial_{x_i} H_q(x^o, y^o)|^2 + |\partial_{y_i} H_q(x^o, y^o)|^2 \right), \quad (2.24)$$

since if $i \in I$ with $\text{dist}(i, \partial I) > d$, we have that $(\partial H_q / \partial z_i)(x^o, y^o) = (\partial H / \partial z_i)(x^o, y^o) = 0$ for $z_i = x_i$ or y_i . For the remaining $2d$ terms denote by P_V the polynomial part of H and note: with again $z_i = x_i$ or y_i ,

$$|\partial_{z_i} H_q(x^o, y^o)|^2 = |\partial_{z_i} P_V(x^o, y^o) - \partial_{z_i} P_V(x^o, y^o; q)|^2,$$

where the notation indicates that coordinates in ∂I are evaluated at either the entries of (x^o, y^o) or the corresponding q . But by pairing entries the above is bounded by a sum of $|q - z_i|^2$ terms with coefficients that are polynomials (of degree at most $2d$) in (x^o, y^o, q) . As we have just shown that $\|(x^o, y^o)\|_\infty$ is uniformly bounded, all these polynomial factors can be further controlled above by a $\rho(q)$ with the claimed properties.

For (2.23) we basically repeat the argument. The key difference being that in the estimate corresponding to (2.24), the sum over $i \in I : \text{dist}(i, \partial I) > d$ on the right hand side does not vanish, but instead produces a multiple of $\sum_{i \in I} \frac{1}{ni^3}$, courtesy Lemma 3. This also explains the restriction of I in this case to $[1, n - d]$. \square

Proof of Proposition 5. Consider first (2.21). The idea is to apply the inequality (2.23) of Lemma 6 to a well-chosen collection of subintervals of I .

Fix an index $i \in I$ and decompose $I \cup \partial I$ into consecutive blocks I_{-m}, \dots, I_m with $\partial I = I_{-m} \cup I_m$ such that $i \in I_0$ and each I_j for $j \neq 0$ is of length d . Denote $J_k = \cup_{|j| \leq k} I_j$ and $\partial J_k = I_{-k-1} \cup I_{k+1}$. Then, as a consequence of (2.23), we have that

$$\|(x^q, y^q) - (x^{\sharp}, y^{\sharp})\|_{2, J_k}^2 \leq c \|(x^q, y^q) - (x^{\sharp}, y^{\sharp})\|_{2, \partial J_k}^2 + c \sum_{i \in J_k} \frac{1}{ni^3}, \quad (2.25)$$

for $k = 0, \dots, m - 1$.

Two remarks are in order. First, a direct application of (2.23) would have the constant c multiplying $\|(x^q, y^q) - (x^{\sharp}, y^{\sharp})\|_{2, \partial J_k}^2$ by a polynomial in the variable (x^q, y^q) appearing in J_k . But (2.23) (or (2.22)) also shows that every $|x^q|$ and $|y^q|$ is bounded by the same polynomial in the variables q . By assumption the q are bounded, and so it is possible to use the same constant c (for all k) throughout (2.25). Second, in case the definition of any I_j places it outside of $\{1, n - d\}$ the corresponding sum is simply taken as empty. This allows the conclusion to extend to one-sided minimizers; when for example I is of the form $[1, L]$ with boundary conditions placed at $[L + 1, L + d]$.

Returning to (2.25), this system is exactly as in the hypothesis of Lemma 7 with

$$a_j = \|(x^q, y^q) - (x^{\sharp}, y^{\sharp})\|_{2, I_{-j} \cup I_j}^2, \quad b_j = \sum_{i \in I_{-j} \cup I_j} \frac{c}{ni^3},$$

and so there is a constant c' for which

$$a_0 \leq c' e^{-m/c'} a_m + c' \max_{j \in [0, m-1]} b_j.$$

This is recognized as (2.21) upon noting: $a_0 \geq (|x_i^q - x_i^{\frac{t}{c}}| + |y_i^q - y_i^{\frac{t}{c}}|)^2$, $a_m \leq d \|q - (x^{\frac{t}{c}}, y^{\frac{t}{c}})\|_{\infty, \partial I}^2$, and $b_j \leq \frac{1}{ni_0^3}$ for all $j \in [0, m-1]$. The proof of (2.20) is identical save that in that case $b_j \equiv 0$. \square

2.3 The limiting mean

The results of the previous subsection yield the following.

Proposition 8. *There is a constant $c = c(V, \beta, a)$ so that*

$$|(x_i^{\frac{t}{c}}, y_i^{\frac{t}{c}}) - (x_i^o, y_i^o)| \leq c' \begin{cases} \frac{1}{\sqrt{n}} & \text{for } i \leq c \log n, \\ \frac{1}{\sqrt{ni^3}} & \text{for } c \log n < i \leq n - c \log n, \\ e^{-(n-i)/c'} & \text{for } n - c \log n < i \leq n, \end{cases} \quad (2.26)$$

with a constant c' depending on c . It follows that

$$\begin{aligned} \sum_{k=\lfloor ns \rfloor}^{\lfloor nt \rfloor} \log \frac{x_k^o}{y_k^o} &= - \left(\frac{a}{2} + \frac{1}{4} \right) \log \frac{\theta(t)}{\theta(s)} + \frac{1}{2} \log \frac{\phi(t)}{\phi(s)} \\ &+ O \left(\frac{1}{ns} + \sum_{k=\lfloor ns \rfloor}^{\lfloor c \log n \rfloor} \frac{1}{\sqrt{k}} + \sum_{k=\lfloor n(1-t) \rfloor}^{\lfloor c \log n \rfloor} e^{-k/c} \right), \end{aligned} \quad (2.27)$$

for all $0 \leq s < t \leq 1$.

For fixed $s < t$ bounded away from 0 and 1, the error term on the righthand side of (2.27) reduces to $O(1/n)$ and we have advertised limiting mean in line one, recall (1.13). The estimate is stated in this more complete form for use in Section 5.

Proof. For (2.26) start with the case that i is a distance $O(\log n)$ away from both 1 and n . Let $I = [i - c \log n, i + \log n]$ and consider the conditional Hamiltonian H_q with $q = (x^o, y^o)$ on ∂I . The conditional minimizer is then the true minimizer through I and we can apply (2.21) of Proposition 5 with $(x^q, y^q) = (x^o, y^o)$. The result is that

$$|(x_i^{\frac{t}{c}}, y_i^{\frac{t}{c}}) - (x_i^o, y_i^o)| \leq c' \max \left(\frac{1}{\sqrt{n(i - c \log n)^3}}, e^{-c'' \log n} \right), \quad (2.28)$$

having used that (x^o, y^o) has bounded entries (proved in the course of establishing Lemma 6). But c'' can be made large with c and by choice $i - c \log n = O(i)$. The other two cases are similar. For example for $i \leq c \log n$, consider $I = [1, c'' \log n]$ with one-sided boundary conditions $= (x^o, y^o)$ on the d -length stretch to the right of $c'' \log n$ (and $c'' \gg c$). Then the boundary component in the analog of (2.28) can be made smaller than any inverse power of n , but the $O(n^{-1/2})$ stemming from $i_0 = 1$ cannot be beat.

Moving to (2.27) we start by noting that

$$\log \frac{x_k^{\frac{1}{2}}}{y_k^{\frac{1}{2}}} = \log \left(\frac{1 + \frac{x^{(1)}(k/n)}{n\phi(k/n)}}{1 + \frac{y^{(1)}(k/n)}{n\phi(k/n)}} \right) = \frac{x^{(1)}(k/n) - y^{(1)}(k/n)}{n\phi(k/n)} + O(k^{-2}),$$

by the estimates of Lemma 4. Summed over $[ns, nt]$ this contributes to $O((ns)^{-1})$ to the advertised error. Next we have that,

$$\sum_{k=\lfloor ns \rfloor}^{\lfloor nt \rfloor} \frac{x^{(1)}(k/n) - y^{(1)}(k/n)}{n\phi(k/n)} = \int_s^t \frac{x^{(1)}(u) - y^{(1)}(u)}{\phi(u)} du + O\left(\frac{1}{ns}\right),$$

where the integral equals the right hand side of the first line in (2.27). The error here follows from the standard Riemann sum bound given that $\left| \frac{d}{du} \left(\frac{x^{(1)}(u) - y^{(1)}(u)}{\phi(u)} \right) \right| = O(u^{-2})$ for small $u > 0$, again by Lemma 4. The remaining overall error term is (the sum of)

$$\log \left(1 + \frac{x_k^o - x_k^{\frac{1}{2}}}{x_k^{\frac{1}{2}}} \right) - \log \left(1 + \frac{y_k^o - y_k^{\frac{1}{2}}}{y_k^{\frac{1}{2}}} \right) = \begin{cases} O(k^{-1/2}) & \text{for } k \leq c \log n, \\ O(k^{-2}) & \text{for } c \log n < k \leq n - c \log n, \\ O(e^{-(n-k)/c}) & \text{for } n - c \log n < k \leq n, \end{cases}$$

Here we have used (2.26) and the fact that $x_k^{\frac{1}{2}}, y_k^{\frac{1}{2}} = O(\sqrt{k/n})$ for small k . The first and third bounds on the right hand side explain the final terms in line two of (2.27). \square

3 Gaussian concentration and approximation

We build up yet more technical machinery. First we establish a sharp form of Gaussian concentration for P about the minimizer (x^o, y^o) of the Hamiltonian H . Along the way we see that similar concentration holds for the conditional distributions of P , or those measures of the form P_q with density proportional $e^{-n\beta H_q(x,y)}$ restricted to the corresponding interval I . Here H_q is the conditional Hamiltonian with boundary conditions q on ∂I introduced in the last section. These estimates are then used to establish approximations of certain P expectations by their Gaussian counterparts.

3.1 Concentration

We re-emphasize that we are assuming $\beta \geq 1$. Our main Gaussian concentration result is the following.

Proposition 9. *There is a constant c' depending on V, β , and a such that for $t > \frac{c'}{\sqrt{n}}$ we have*

$$P(|X_k - x_k^o| + |Y_k - y_k^o| > t) \leq ce^{-nt^2/c} \quad (3.1)$$

for any $k \in [1, n]$ and c depending on c' .

This serves as a refinement of the Brascamp-Lieb type inequality proved as Lemma 8.1 of [17]:

Lemma 10. *There is a constant $c = c(V, \beta, a)$ so that $\|(X, Y) - (x^o, y^o)\|_2$ is stochastically dominated by $\|G\|_2$ where G is the Gaussian vector on \mathbb{R}^{2n-1} with density proportional to $e^{-cn\|g\|_2^2/2}$. Additionally, under any P_q we have that $\|(X, Y) - (x^q, y^q)\|_{2,I}$ is similarly dominated by the norm of a Gaussian vector in dimension $2|I|$ with entry variance $(cn)^{-1}$.*

Note that [17] states the above in a far more general way. What is important for the present application is that the convexity constant for any conditional H_q can be bounded below by that of H . Lemma 10 may be iterated to produce increasingly better tail estimates on local scales (or shorter stretches of indices). While this sufficed for the soft edge problem in [17], for the control required below it is more efficient to take a different approach.

Here we rely on the fact that, with $\beta \geq 1$, P satisfies a Logarithmic Sobolev Inequality (see for example [2]). Given that, “Herbst’s argument” (see Theorem 5.3 of [16]) yields:

Lemma 11. *There is a constant $c = c(V, \beta, a)$ such that*

$$P(|F(X, Y) - EF(X, Y)| > t) \leq 2e^{-ct^2/2\|F\|_{Lip}^2},$$

for any Lipschitz function $F : \mathbb{R}_+^{2n-1} \mapsto \mathbb{R}$.

Applied to $F(x, y) = x_k$ or y_k this produces an inequality of the form (3.1) for all $t > 0$, though centered at the mean rather than at the minimizer. Proposition 9 then follows from the next estimate, which actually makes essential use of the old Gaussian concentration result Lemma 10.

Lemma 12. *It holds that*

$$E\left[|X_k - x_k^o| + |Y_k - y_k^o|\right] \leq \frac{c}{\sqrt{n}},$$

for any $k \in [1, n]$ and $c = c(V, \beta, a)$.

Proof. Throughout we use the shorthand Z (or z) to denote the pair of variables (X, Y) (or (x, y)).

To start we fix an interval $D = \{k \in [\ell - d/2, \ell + d/2]\}$, and for a choice of m (to be determined) let I be the interval made up of D and the (at most) m indices to the left/right. As usual ∂I will denote the d indices to the left and/or right of I . If any part of D , I , or ∂I falls outside of $[1, n]$, it is truncated in the obvious way or viewed as empty. Then, with E_q the conditional expectation with respect to the variables $q \in \partial I$ we write

$$E \max_{k \in D} |Z_k - z_k^o| \leq EE_q \max_{k \in D} |Z_k - z_k^q| + EE_q \max_{k \in D} |z_k^o - z_k^q|. \quad (3.2)$$

Here z^q denotes the conditional minimizer of H_q on I (with boundary conditions q). By passing the randomness onto the variables q (for which we continue to use lower case) we will be able to iterate this inequality.

Further bounding (3.2) above we have that

$$EE_q \max_{k \in D} |Z_k - z_k^q| \leq EE_q \|Z - z^q\|_{2,I} \leq c \sqrt{\frac{m}{n}} \quad (3.3)$$

by Lemma 10. And by Proposition 5 we also have the bound: with $Q_{\partial I}$ the event that $|q - z^0|_{\infty, \partial I} < b$,

$$\max_{k \in D} |z_k^o - z_k^q| \mathbf{1}_{Q_{\partial I}} \leq c' e^{-m/c} \max_{i \in \partial I} |z_i^o - q_i|, \quad (3.4)$$

for a $c = c'(b)$. And using Lemma 6 on the complement of $Q_{\partial I}$:

$$EE_q \max_{k \in D} |z_k^o - z_k^q| \leq c' e^{-m/c} E \left[\max_{i \in \partial I} |z_i^o - q_i| \right] + E \left[\rho(q) \|z^o - q\|_{2, \partial I}, Q_{\partial I}^c \right].$$

By Holder's inequality and another application of Proposition 10 we can control the second term on the right hand side by a constant multiple of $P(Q_{\partial I}^c) \leq c' e^{-n/c'}$ with a new $c' = c'(b)$.

Adjusting constants and substituting this last estimate along with (3.3) into (3.2) gives

$$E \max_{k \in D} |Z_k - z_k^o| \leq c \sqrt{\frac{m}{n}} + c e^{-m/c} E \left[\max_{i \in \partial I} |z_i^o - q_i| \right] + c e^{-n/c}. \quad (3.5)$$

At this point we can choose m large enough (but independent of n) so that $c e^{-m/c} < 1/4$ and then for large enough n absorb the final term on the right hand side into the first. Then (3.5) may be schematized as in:

$$a_k \leq 2c \sqrt{\frac{m}{n}} + \frac{1}{4}(a_{k-m} + a_{k+m}). \quad (3.6)$$

Here a_k is $E \max_{i \in D} |Z_i - z_i^o|$ for whatever interval D centered at k with $a_\ell = 0$ for $\ell \leq 0$ or $\ell \geq n$. In this interpretation, the q in the expectation on the right hand side of (3.5) stands in for the corresponding (random) Z variable while ∂I serves as a shifted copy of D .

The claim is that (3.6) (along with its corresponding side conditions) implies all the $a_k = a_k(n)$ are bounded by a constant multiple of $n^{-1/2}$. After a scaling the problem can be summarized thus: Given an array $b_k = b_k(L)$, for $k = 0, \dots, L$ which is nonnegative, finite and satisfies (for each L),

$$b_k \leq 1 + \frac{1}{4}(b_{k-1} + b_{k+1}), \quad b_0 = b_L = 0, \quad (3.7)$$

there is a constant which bounds all b_k independently of L . This can be seen by contradiction: if for some j we have that say $b_j \geq 4$, any such solution must grow exponentially to either the left or right of j . But this would violate the Dirichlet boundary conditions imposed at $j = 0$ or $j = L$. \square

3.2 Laplace estimates

Along with $P_q(dx, dy) = \frac{1}{Z_q} e^{-n\beta H_q(x, y)} dx dy$ introduce the natural Gaussian measure approximating P_q over the same interval I :

$$\nu_{I, q}(dx, dy) = \frac{1}{Z'_q} \exp \left(-\frac{n\beta}{2} \left\langle (x - x^q, y - y^q), \mathcal{H}_q(x - x^q, y - y^q) \right\rangle \right) dx dy, \quad (3.8)$$

where \mathcal{H}_q denotes the Hessian of H_q evaluated at (x^q, y^q) . We also bring in the mixture of ν_q over boundary conditions in “typical” position, defined by

$$\int F(x, y) \mu_{I, c}(dx, dy) = E \left[\int F(x, y) \nu_{I, q}(dx, dy), \|q - (x^o, y^o)\|_{\infty, \partial I} \leq c\delta_n \right] \quad (3.9)$$

with $\delta_n = \sqrt{\frac{\log n}{n}}$.

To determine the statistics of the field for bulk indices, we have the following estimate which relates the P -expectation of certain polynomial test functions to those of averaged Gaussians.

Proposition 13. *Fix a small $\delta > 0$ and let $I \subset [\delta n, n]$. Denote by K the interval made up of I along with the (at most) $c \log n$ indices to its left and right. Let $F_I(x', y')$ be a nonnegative polynomial with bounded coefficients and of bounded degree in the variables $(x_i - x'_i)$ and $(y_i - y'_i)$ for $i \in I$ and prescribed centerings (x', y') . Then, there exist constants c and c' (which depend on V, β, a and the degree of F_I) such that*

$$E[F_I(x^o, y^o)] = \left(\int F_I(x^q, y^q) \mu_{K, c'}(dx, dy) + O(n^{-2}) \right) \left(1 + O \left(\frac{|K|(\log n)^{3/2}}{\sqrt{n}} \right) \right). \quad (3.10)$$

Of course, c and c' figure into the implied constants in the error terms and the estimate (3.10) presumes that $|K|(\log n)^{3/2} n^{-1/2} = o(1)$.

As a consequence of Proposition 13 the limiting variance of the field $k \mapsto (X_k, Y_k)$ will be determined through:

Corollary 14. *Now let $I = [i_o, i_1]$ of length at most $n^{1/4-}$ and supported in $[\delta n, n - 2c \log n]$ for fixed $\delta > 0$ and large $c = c(V, \beta, a)$. With $K = [i_o - c \log n, i_1 + c \log n]$, denote by P_q be the conditional measure on K with boundary conditions q satisfying $\|q - (x^o, y^o)\|_{\partial K, \infty} \leq c' \sqrt{\frac{\log n}{n}}$ (for a large $c' = c'(V, \beta, a)$). Then*

$$E_q \left[\left(\sum_{i \in I} (X_i - x_i^o) - (Y_i - y_i^o) \right)^2 \right] = \phi^2(i_o/n) \frac{\theta'(i_o/n)}{\theta(i_o/n)} \frac{(i_1 - i_o)}{\beta n} + O \left(\frac{(\log n)^2}{n} \right), \quad (3.11)$$

where the implied constant in the error term depends only on $\delta, V, \beta, a, c, c'$. In particular, the same estimate holds with E in place of E_q .

Last, we will require the more particular control for indices down to $O(\log n)$ away from the singularity.

Corollary 15. *It holds that*

$$E[(X_k - x_k^o) - (Y_k - y_k^o)] = O\left(\frac{(\log k)^5}{\sqrt{nk}}\right), \quad (3.12)$$

and

$$E[(X_k - x_k^o)^2 - (Y_k - y_k^o)^2] = O\left(\frac{(\log k)^{5/2}}{n\sqrt{k}}\right), \quad (3.13)$$

uniformly for $k \in [c \log n, n - c \log n]$ with $c = c(V, \beta, a)$ sufficiently large.

While Corollary 14 is a direct calculation based on Proposition 13, the proof of Corollary 15 entails that a higher order expansion be made than that behind the estimate (3.10). (The unattractive log factors in (3.12) and (3.13) could be improved by yet a higher order expansion, but the above suffices for what we will need.)

Proof of Proposition 13. To be concrete, we will assume that there are constants p, q so that $F_I \leq q + q\|(x, y) - (x^o, y^o)\|_{I,2}^p$. It will be clear in the course of the argument that other choices of (bounded coefficient and bounded degree) polynomial F_I will only alter the choices of c' and c made along the way.

With F_I as specified and $Q_{\partial K} = \{ \|q - (x^o, y^o)\|_{\partial K, \infty} \leq c' \sqrt{\frac{\log n}{n}} \}$, we first claim that by choice of c' and c :

$$\begin{aligned} E[F_I(x^o, y^o)] &= E[E_q[F_I(x^o, y^o)], Q_{\partial K}] + O(n^{-2}) \\ &= E[E_q[F_I(x^q, y^q)], Q_{\partial K}] + O(n^{-2}), \end{aligned} \quad (3.14)$$

where (x^q, y^q) refers to the minimizer of the corresponding H_q on the larger interval K . (Note: if the right edge of support of I is less than $c \log n$ away from n , we are considering a one-sided minimization with boundary conditions placed to the left of K .) For line one, Lemma 10 shows $EF_I^2 = O(1)$ while for c' large enough $P(Q_{\partial K}^c) = O(n^{-4})$ by Proposition 9 – indeed the (-4) may be replaced by any negative power by choice of c' . Then apply Cauchy-Schwartz. For line two we assume that c is large enough depending on c' , so that Proposition 5 provides: for any $i \in I$, $|x_i^q - x_i^o|, |y_i^q - y_i^o| \leq c''n^{-4}$ with a c'' uniform over $q \in Q_{\partial K}$. A second application of Cauchy-Schwartz using Lemma 10 to control the E_q expectation of powers of $F_I(x^q, y^q)$ produces the estimate.

At the expense of another $O(n^{-2})$ error, we can now further restrict the inner E_q expectation in (3.14) to the event Q_K on which $|X_i - x_i^q|, |Y_i - y_i^q| \leq c''' \sqrt{\frac{\log n}{n}}$ for all $i \in K$ for some c''' . This is a repetition of the argument employed in the first estimate of (3.14) coupled

with the fact that Proposition 5 gives that $|x_i^q - x_i^o| \vee |y_i^q - y_i^o|$ is $O(\sqrt{\log n/n})$ throughout K (with the sharper estimate used just above holding on I)

Now we are in position to approximate P_q by the Gaussian measure $\nu_q = \nu_{K,q}$ introduced in (3.8). If necessary we can adjust Z_q (the P_q normalizer) so that $H_q(x^q, y^q) = 0$. Then, with

$$H_q^{(3)}(x, y) = H_q(x, y) - \left\langle (x - x^q, y - y^q), \frac{1}{2} \mathcal{H}_q(x - x^q, y - y^q) \right\rangle, \quad (3.15)$$

Taylor's formula gives that

$$|H_q^{(3)}(x, y)| \leq \sup_{t \in [0,1]} \sum_{k \in K} \left(\rho_k(t) + \frac{4k}{n|x_k(t)|^3} + \frac{4k}{n|y_k(t)|^3} \right) |(x_k, y_k) - (x_k^q, y_k^q)|^3. \quad (3.16)$$

Here $\rho_k = \rho_{V,k}$ indicates a (positive) polynomial of fixed degree in the variables $(x_i(t), y_i(t))$ for $i \in [k-d, k+d]$, while $(x(t), y(t))$ draws out the line between (x^q, y^q) and (x, y) . We use the fact that $\text{tr}V(BB^T)$ is finite-range, so there are fixed number of mixed third-partial derivative involving any index $k \in K$ stemming from the polynomial of H_q . For the factors corresponding to the third derivatives of the log terms in H_q , note that any $k \in K$ under consideration is large enough so that $4k \geq 2(k + |a| + 1/\beta)$.

Further, with the left endpoint of K at least $\delta n/2$, the results of the Section 2 gives that: restricted to Q_K , (x^q, y^q) , and so also $(x(t), y(t))$, are bounded above and below independently of n or $q \in Q_{\partial K}$. Therefore,

$$n\beta |H_q^{(3)}(x, y)| \mathbf{1}_{Q_K} \leq \gamma |K| \left(\frac{(\log n)^{3/2}}{n^{1/2}} \right) \quad (3.17)$$

with γ depending only on the parameters δ, c, V, β, a . This allows the conclusion that

$$E_q[F_I \mathbf{1}_{Q_K}] = \int_{Q_K} F_I(x^q, y^q) \nu_q(dx, dy) \left(1 + O\left(\frac{(\log n)^{3/2} |K|}{\sqrt{n}} \right) \right), \quad (3.18)$$

which is effectively the claim. Here we have used that, with η the right hand side of (3.17),

$$e^{-\eta} \nu_q(Q_K) \leq \frac{Z_q}{Z'_q} \leq e^\eta \frac{1}{P_q(Q_K)}. \quad (3.19)$$

The Logarithmic Sobolev Inequality for Gaussian measures gives that $\nu_q(Q_K)$ is the same order as $P_q(Q_K)$ (the measures were built to have the same convexity constant). That is, with each of these factors a negative power of n , the upper and lower bounds in (3.19) are controlled by $e^{\pm\eta} = 1 + O(\eta)$ with $\eta = o(1)$. \square

Proof of Corollary 14. Denoting by $S_I(x^o, y^o)$ the squared sum within the expectation of (3.11), the (beginning of) the proof of Proposition 13 yields $E_q[S_I(x^o, y^o)] = E_q[S_I(x^q, y^q) \mathbf{1}_{Q_K}]$

$+O(n^{-2})$. Again, Q_K is the event that $\|(x, y) - (x^q, y^q)\|_{\infty, K}$ is less than $c' \sqrt{\frac{\log n}{n}}$ (for choice of c'). Continuing, the same proposition gives that

$$E_q[S_I(x^q, y^q)\mathbf{1}_{Q_K}] = \left(\int S_I(x^q, y^q) \nu_q(dx, dy) + O(n^{-2}) \right) \left(1 + O\left(\frac{(\log n)^{3/2}|K|}{\sqrt{n}} \right) \right), \quad (3.20)$$

and we will show that the advertised appraisal (3.11) holds for the remaining ν_q integral. This is enough since the multiplicative error in (3.20), can be combined with the $O(n^{-1}|K|)$ leading order term in (3.11) to be absorbed into a second $o((\log n)^2/n)$ additive error. Note that the factor $\phi^2(i_0/n) \frac{\theta'(i_0/n)}{\theta(i_0/n)}$ is of order one.

Next compute:

$$\int S_I(x^q, y^q) \nu_q(dx, dy) = \frac{1}{n\beta} w^T \mathcal{H}_q^{-1} w, \quad (3.21)$$

where the right hand side is read as follows. Indexing the integral (and so \mathcal{H}_q) according to $(x_{k_0}, y_{k_0}, x_{k_0+1}, y_{k_0+1}, \dots, x_{k_1}, y_{k_1})$ where $[k_0, k_1] = K$, the $(2|K|)$ -vector w has entries $(-1)^i$ for indices corresponding to the coordinates in $I \subset K$ and is otherwise zero. To estimate $w^T \mathcal{H}_q^{-1} w$ we approximate \mathcal{H}_q^{-1} by its “coarse” version and find an exact eigenvalue problem. Here is where we will use the assumption that I is supported $O(\log n)$ away from n .

Recall the coarse Hamiltonian on K at position $t = \frac{(k_0+k_1)}{2n} = \frac{(i_0+i_1)}{2n}$ introduced in (2.8),

$$H_K(x, y) = \text{tr} V(CC^T) - t \sum_{k \in K} \log(x_k y_k).$$

Here C is the $m \times m$ circulant version of the matrix $B(x, y)$. Then, with \mathcal{H}_* the Hessian of H_K at its minimizer $x_k = y_k = \phi(t)$ for $k \in K$, the needed fact is:

$$v^T (\mathcal{H}_*)^{-1} v = |K| \phi(t)^2 \frac{\theta'(t)}{\theta(t)}, \quad (3.22)$$

for the vector v with entries $v_k = (-1)^k, k = 1, \dots, 2|K|$. In particular, this is approximation of the $w^T \mathcal{H}_q^{-1} w$ appearing in (3.21). Comparing this to the statement (3.11), note that there we sample ϕ and θ at the initial point i_0/n rather than the midpoint $(i_0 + i_1)/n$ – the error in going back in forth between this two is easily seen to be of sufficiently lower order.

To see (3.22), note that \mathcal{H}_K is circulant Toeplitz, and hence has v as an eigenvector. With λ the corresponding eigenvalue: with $(z_1, z_2, z_3, z_4, \dots) = (x_{k_0}, y_{k_0}, x_{k_0+1}, y_{k_0+1}, \dots)$,

$$\lambda = \left(\sum_{1 \leq k, \ell \leq 2|K|} (-1)^{k+\ell} \frac{\partial^2 H_K}{\partial z_k \partial z_\ell} \right) \Big|_{z_i = \phi(t), i=1, \dots, 2|K|} = \frac{1}{2} \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right)^2 H_t(x, y) \Big|_{x=y=\phi(t)},$$

where we bring back our initial definition of the coarse Hamiltonian H_t from (2.1). Using that formula we find that

$$\lambda = \frac{t}{\phi(t)^2} + \sum_{m=1}^d g_m \frac{m}{2m-1} \binom{2m}{m} \phi(t)^{2m-2} = \frac{1}{\phi(t)} \int_0^t \frac{ds}{\phi(s)} = \frac{2\theta(t)}{\phi^2(t)\theta'(t)}.$$

The middle equality follows from the definition (1.7) for ϕ upon multiplying both sides of that identity by $\phi'(t)\phi^{-2}(t)$ and integrating by parts.

It remains to show that

$$|w^T \mathcal{H}_q^{-1} w - v^T (\mathcal{H}_*)^{-1} v| = O((\log n)^2). \quad (3.23)$$

Consider the matrix \mathcal{H} formed by setting all entries in the $O(1)$ blocks in the upper right and lower left corners of \mathcal{H}_* to zero. Clearly we have that $|v^T \mathcal{H}^{-1} v - v^T (\mathcal{H}_*)^{-1} v| = O(1)$. On the other hand \mathcal{H}_q and \mathcal{H} are also nearly the same. They are $2d$ -banded with corresponding entries built from the same functions – except along the diagonal – evaluated at either (x^q, y^q) or $(\phi(t), \phi(t))$. Along the diagonal the functional entries differ only in the coefficients of the terms corresponding to the second derivatives of the logarithm in H_q or $H_{K,t}$: there one must compare $t\phi^{-2}(t)$ to $(k + a - 1/\beta)n^{-1}(x_k^q)^{-2}$ or $(k - 1/\beta)n^{-1}(y_k^q)^{-2}$ for any $k \in K$. But these coefficients (t and $(k + a - 1/\beta)/n$ or $(k - 1/\beta)/n$) are no more than $O(|K|n^{-1}) = O(n^{-3/4})$ apart. Further, restricted to Q_K the values of (x^q, y^q) and $(\phi(t), \phi(t))$ are no more than $O(\sqrt{\frac{\log n}{n}})$ apart (and are uniformly bounded below). It follows that

$$\|\mathcal{H}_q^{-1} - \mathcal{H}^{-1}\| = \|\mathcal{H}_q^{-1}\| \|\mathcal{H}^{-1}\| \|\mathcal{H}_q - \mathcal{H}\| \leq c'' \sqrt{\frac{\log n}{n}}. \quad (3.24)$$

Convexity of H_q provides a constant upper bound on $\|\mathcal{H}_q^{-1}\|$ and $\|\mathcal{H}^{-1}\|$. Then, the previous remarks along with the Gershgorin circle theorem yield $\|\mathcal{H}_q - \mathcal{H}\| = O(\sqrt{\frac{\log n}{n}})$. It remains to show that

$$|(w - v)^T \mathcal{H} v| = O((\log n)^2),$$

and likewise for $(w - v)^T \mathcal{H}^{-1} w$. Noting that $(w - v)$ has only $O(\log n)$ non-zero entries, this is a consequence of the banded Toeplitz nature of \mathcal{H} which produces exponential decay in the entries of \mathcal{H}^{-1} away from the diagonal. \square

Proof of Corollary 15. The idea is similar to that behind Proposition 13, though now for each $k \in [c \log n, n - c \log n]$ we let K be the interval of length $c \log k$ centered at k for the constant c to be chosen momentarily.

Let again q denote the coordinates in ∂K , but now let $Q_{\partial K}$ be the event that $\|q - (x^o, y^o)\|_{\partial K, \infty}$ is less than $c' \sqrt{\frac{\log k}{n}}$. By choice of c' and Proposition 9, $P(Q_{\partial K}) = 1 - O(k^{-4})$ since $k \gg 1$. The same proposition gives that $E[(X_k - x_k^o)^{2p} + (Y_k - y_k^o)^{2p}] = O(n^{-p})$. Both estimates are uniform in k . And so, by the Cauchy-Schwartz (and Jensen's) inequality $E[E_q | X_k - x_k^o|^p, Q_{\partial K}^c] = O(n^{-p} k^{-2})$, and likewise in the y -variable. Next, for any given q in a given $Q_{\partial K}$, we can select the $c = c(\beta, V, a)$ so that Proposition 5 gives $|x_k^o - x_k^q| + |y_k^o - y_k^q| = O(e^{-c''(c) \log k} (\log k/n)^{1/2}) = O(k^{-2} n^{-1/2})$. The conclusion is that: for $p = 1$ or 2 ,

$$E[(X_k - x_k^o)^p - (Y_k - y_k^o)^p] = E\left[E_q\left[\Delta_p(X_k, Y_k) \mathbf{1}_{Q_K}\right] \mathbf{1}_{Q_{\partial K}}\right] + O(n^{-p/2} k^{-2}), \quad (3.25)$$

uniformly in k . Here we have made the definition,

$$\Delta_p(x_k, y_k) = (x_k - x_k^q)^p - (y_k - y_k^q)^p,$$

and $Q_K = Q_K(q)$ is the event $\{\|(x, y) - (x^q, y^q)\|_{K, \infty} \leq c' \sqrt{\frac{\log k}{n}}\}$, for a possibly adjusted c' . That we can restrict the E_q integral in (3.25) to Q_K with the stated level of error, follows from the same argument used at the analogous step in the proof of Proposition 13.

Turning to an estimate on $E_q[\Delta_p(X_k, Y_k) \mathbf{1}_{Q_K}]$, we start with the case $p = 2$. Under the approximating Gaussian measure $\nu_q = \nu_{K, q}$ we have that

$$\int_{Q_K} \Delta_2(x_k, y_k) \nu_q(dx, dy) = \frac{1}{n\beta} \left((\mathcal{H}_q^{-1})_{2k-1, 2k-1} - (\mathcal{H}_q^{-1})_{2k, 2k} \right) + O\left(\frac{1}{nk^2}\right). \quad (3.26)$$

The first term is an exact Gaussian computation, after removing the restriction to Q_K . The error term uses that by choice of c' it holds $\nu_q(Q_K) = 1 - O(k^{-4})$. With $K = [k_0, k_1]$ the indices of the ν_q integral and the matrix \mathcal{H}_q are indexed $x_{k_0}, y_{k_0}, x_{k_0+1}, y_{k_0+1}, \dots$, as before.

Next recall that the proof of Corollary 14 introduces a banded Toeplitz approximate \mathcal{H} to \mathcal{H}_q (\mathcal{H} is the Hessian of the coarse Hamiltonian H_K on K , with the corner entries which make the latter circulant removed), and would like to replace the appearances of \mathcal{H}_q in (3.26) with this approximate. Since $(\mathcal{H}^{-1})_{ii} = (\mathcal{H}^{-1})_{jj}$ for all $i, j \in K$ and, by convexity, $\|\mathcal{H}_q^{-1} - \mathcal{H}^{-1}\|$ is controlled by a constant multiple of $\|\mathcal{H}_q - \mathcal{H}\|$, this norm must now be estimated for K possibly within $\log n$ of the singularity.

In the current setting we have that: with $t = k/n$ and so $\phi(t)$ the common variable where the entries of \mathcal{H} are evaluated, $|x_i^q - \phi(t)| + |y_i^q - \phi(t)| = O(\sqrt{\log k/n})$. Hence the difference between any off diagonal of \mathcal{H} and \mathcal{H}_q are also controlled by $O(\sqrt{\log k/n})$. The more delicate issue is now the diagonals where one has to consider the absolute differences $|\phi(t)^{-2} - (x_i^q)^{-2}|$ or $|\phi(t) - (y_i^q)^{-2}|$ which are $O((n/k)^{3/2} \times \sqrt{\log k/n})$. Here we use that in general we have that $\phi(t) \geq \delta \sqrt{k/n}$, and so the given x_i^q and y_i^q for $i \in K$ satisfy the same lower bound. But since any of these diagonal components are multiplied by coefficients which are $O(k/n)$ throughout K , the corresponding entry differences are actually $O(\sqrt{\log k/k})$ and we have that

$$\|\mathcal{H}_q - \mathcal{H}\| = O\left(\sqrt{\frac{\log k}{k}}\right), \quad (3.27)$$

compare (3.24). Therefore, (3.26) can be continued as in

$$\int_{Q_K} \Delta_2(x_k, y_k) \nu_q(dx, dy) = O\left(\frac{\sqrt{\log k}}{n\sqrt{k}}\right). \quad (3.28)$$

To finish, write

$$E_q[\Delta_2(X_k, Y_k), Q_K] = \frac{Z'_q}{Z_q} \int_{Q_K} \Delta_2 d\nu_q + \frac{Z'_q}{Z_q} \int_{Q_K} \Delta_2 (e^{n\beta H_q^{(3)}} - 1) d\nu_q, \quad (3.29)$$

where once more Z'_q and Z_q denote the normalizers for ν_q and P_q , respectively. Now recalling (3.16) from the proof of Proposition 13, the estimate (3.17) can be replaced by

$$n\beta|H_q^{(3)}(x, y)|\mathbf{1}_{Q_K} \leq \gamma \left(\frac{(\log k)^{5/2}}{k^{1/2}} \right), \quad (3.30)$$

for another constant $\gamma = \gamma(V, \beta, a, c, c')$. Here we have used the current definition of Q_K which restricts $\|(x^q, y^q) - (x, y)\|_{K, \infty}^3$ to $O((\log k/k)^{3/2})$, that now $|K| = c \log k$, and a worse case upper bound of $O((n/k)^{3/2})$ on any of the $x_i(t)^{-3}$ or $y_i(t)^{-3}$ for $i \in K$ (recall that these are the interpolants from (x_i^q, y_i^q) to $(x_i, y_i) \in Q_K$). Next, since the right hand side of (3.30) is $o(1)$ for $k \geq c \log n$, (3.19) shows the ratios Z'_q/Z_q are bounded above and below by constants only depending on c, c' and V, β, a . Finally then, using that $|e^\zeta - 1| \leq 2|\zeta|$ for $|\zeta| \leq 1$ (applied to $\zeta = n\beta H_q^{(3)}$ restricted to Q_K) and $\int |\Delta_2| d\nu_q = O(n^{-1})$, we find for the second term in (3.29) that

$$\int_{Q_K} \Delta_2(e^{n\beta H_q^{(3)}} - 1) d\nu_q = O\left(\frac{(\log k)^{5/2}}{nk^{1/2}}\right). \quad (3.31)$$

This is the estimate reported in (3.13).

For the difference of the means, one has to consider an additional order. We now write,

$$E_q[\Delta_1(X_k, Y_k), Q_k] = \frac{Z'_q}{Z_q} \int_{Q_K} \Delta_1(1 + nH_q^{(3)}) d\nu_q + \frac{Z'_q}{Z_q} \int_{Q_K} \Delta_1(e^{n\beta H_q^{(3)}} - 1 - nH_q^{(3)}) d\nu_q,$$

for which we readily have the following:

$$\begin{aligned} \int_{Q_K} \Delta_1 d\nu_q &= O(n^{-1/2}k^{-2}), \\ \int_{Q_K} \Delta_1(e^{n\beta H_q^{(3)}} - 1 - nH_q^{(3)}) d\nu_q &= O\left(\frac{(\log k)^5}{n^{1/2}k}\right). \end{aligned} \quad (3.32)$$

The first of these is due: $\int \Delta_1 d\nu_q = 0$, $\int |\Delta_1|^2 d\nu_q = O(n^{-1})$, while c' can be chosen so that $\nu_q(Q_K^c) = O(k^{-4})$. (The displayed estimate follows from applying Cauchy-Schwartz to integral over Q_K^c). The second is similar to (3.31): now $\int |\Delta_1| d\nu_q = O(n^{-1/2})$ while $|e^{n\beta H_q^{(3)}} - 1 - nH_q^{(3)}|$ on Q_K is controlled by the square of the right hand side of (3.30).

As we have noted, Z'_q/Z_q is of constant order (uniformly for all K and choices of “good” boundary conditions q), and so it remains to consider $n \int_{Q_K} \Delta_1 H_q^{(3)} d\nu_q$. For this we first schematize $H_q^{(3)}$ as in

$$H_q^{(3)}(x, y) = P_V(x, y) + \sum_{i \in K} \left(a_i(x_i - x_i^q)^3 + b_i(y_i - y_i^q)^3 \right). \quad (3.33)$$

Here $P_{V,q}$ represents the appropriate sum of third derivatives of the potential term, while with

$$a_i = 2x_i(t)^{-3}(i/n + a/n - 1/n\beta), \quad b_i = 2y_i(t)^{-3}(i/n - 1/n\beta),$$

the sum over centered cubics corresponds to the (third derivatives of the) logarithmic terms of H_q . Since a_i and b_i are complicated functions of (x, y) , to perform the desired integral we first note: on Q_K ,

$$|a_i - c_i| + |b_i - c_i| = O\left(\frac{\sqrt{n \log k}}{k}\right), \quad \text{with } c_i = \frac{2i}{n\phi^3(i/n)}.$$

Hence, if we replace all appearances of a_i and b_i in (3.33) by c_i , we make an $O(\log k \times \frac{\sqrt{n \log k}}{k} \times \frac{(\log k)^{3/2}}{n^{3/2}}) = O(\frac{(\log k)^2}{nk})$ sup-norm error (granted we working on Q_K) to $H_q^{(3)}$, and so a $O(\frac{(\log k)^2}{\sqrt{nk}})$ error in any estimate of $n \int_{Q_K} \Delta_1 H_q^{(3)} d\nu_q$.

Similar preprocessing is required for the integral involving P_V . However, that integral will clearly be subdominant compared with that over the second term in (3.33) since both a_i and b_i are as large as $O(\sqrt{n/i})$. We will therefore only detail how to deal with this term. After making the substitution just described, we have the evaluation:

$$\begin{aligned} \int \Delta_1(x_k, y_k) \left(n \sum_{i \in K} c_i ((x_i - x_i^q)^3 + (y_i - y_i^q)) \right) d\nu_q \\ = \frac{1}{n} \sum_{i \in K} (3c_i) \left[(\mathcal{H}_q^{-1})_{2i-1, 2i-1} \left((\mathcal{H}_q^{-1})_{2k-1, 2i-1} - (\mathcal{H}_q^{-1})_{2k, 2i-1} \right) \right. \\ \left. + (\mathcal{H}_q^{-1})_{2i, 2i} \left((\mathcal{H}_q^{-1})_{2k-1, 2i} - (\mathcal{H}_q^{-1})_{2k, 2i} \right) \right]. \end{aligned} \quad (3.34)$$

It is by now understood that we can go from this full-space integral to that restricted to Q_K making further subdominant errors. Things are at long last wrapped in the same way that (3.26) was treated. First observe that, if we could replace \mathcal{H}_q with its approximate \mathcal{H} throughout (3.34), the quantity within the square brackets vanishes on account that \mathcal{H} is Toeplitz. Since again all entries of \mathcal{H}_q^{-1} and \mathcal{H}^{-1} are uniformly bounded, a computation using (3.27) shows that the error incurred in making that substitution in (3.34) is $O(\frac{(\log k)^{7/2}}{\sqrt{nk}})$. As this lies under the larger of the error estimates in (3.32) – which is what is reported in (3.12) – the proof is finished. \square

4 Central limit theorem

Here we complete the identification of the limit of the K_n kernel by proving:

Proposition 16. *As $n \rightarrow \infty$,*

$$\sum_{k=\lfloor nt \rfloor}^n \log \frac{X_k/x_k^o}{Y_k/y_k^o} \Rightarrow \frac{1}{\sqrt{\beta}} \int_{\theta(t)}^1 \frac{db_u}{\sqrt{u}},$$

in the Skorohod topology on $(0, 1]$.

Recall (1.14). Note that Gaussian concentration plus the formulas for the minimizer developed in the last two sections already give that $X_{[nt]} \Rightarrow \phi(t)$ as processes on $[\delta, 1]$ for any $\delta > 0$.

4.1 Linearizing

As a first step we have the following.

Lemma 17. *For Proposition 16 it is sufficient to show that*

$$\sum_{k=[nt]}^n \frac{(X_k - x_k^o) - (Y_k - y_k^o)}{\phi(k/n)} \Rightarrow \frac{1}{\sqrt{\beta}} \int_{\theta(t)}^1 \frac{db_u}{\sqrt{u}},$$

in the Skorohod topology on $(0, 1]$.

Proof. We fix a (small) $\delta > 0$, and show the claim for all processes restricted to $t \in [\delta, 1]$. Afterwards it will be clear that the choice of δ is arbitrary.

Again denote

$$Q = \left\{ |X_k - x_k^o|, |Y_k - y_k^o| \leq c \sqrt{\frac{\log n}{n}} \text{ for all } k \in [1, n] \right\},$$

with c chosen so that $P(Q^c) \leq n^{-4}$ for all n large enough (Proposition 9). Certainly it is enough to work with the process $\mathbf{1}_Q \sum_{k=[nt]}^n \log \frac{X_k/x_k^o}{Y_k/y_k^o}$. At the same time,

$$\mathbf{1}_Q \left[\log \frac{X_k/x_k^o}{Y_k/y_k^o} - \left(\frac{(X_k - x_k^o) - (Y_k - y_k^o)}{\phi(k/n)} \right) + \left(\frac{(X_k - x_k^o)^2 - (Y_k - y_k^o)^2}{2\phi(k/n)^2} \right) \right] = O(n^{-(3/2-\epsilon)}),$$

uniformly for $k \in [n\delta, n]$ with probability one. Here ϵ can be chosen as small as one likes subject to the implied constant on the right hand side depending on ϵ . This follows as $|\log(1+t) - t + t^2/2| \leq |t|^3$ for all $|t| \leq 1/2$ while, with $(Z, z) = (X, x)$ or (Y, y) : on Q all $|Z_k - z_k^o|^p$ are $O(n^{-(p/2-\epsilon)})$, z_k^o is uniformly bounded below for $k > \delta n$, and $|\frac{1}{z_k} - \frac{1}{\phi(k/n)}| = O(n^{-1})$ throughout the same range of indices. For the last two facts see Proposition 8. So, with the left hand side of the above display denoted by $\eta_{k,n}$ we have that $t \mapsto \sum_{[nt]}^n \eta_{k,n}$ converges to the zero process.

Consider next

$$\zeta_m^n(X, Y) = \mathbf{1}_Q \sum_{k=m}^n \frac{(X_k - x_k^o)^2 - (Y_k - y_k^o)^2}{\phi(k/n)^2}.$$

The proof will be finished by showing that $\max_{m \in [\delta n, n]} |\zeta_m^n|$ goes to zero with probability one. We actually take the approximation

$$\hat{\zeta}_m^n(X, Y) = \sum_{k=m}^n \frac{\psi(X_k - x_k^o) - \psi(Y_k - y_k^o)}{\phi(k/n)^2},$$

for $\psi(z) = z^2$ for $|z| \leq c\sqrt{\frac{\log n}{n}}$ outside of which ψ is taken to be constant. Obviously, $\hat{\zeta}_m^n$ and ζ_m^n agree on Q , while as a map from $x_m, y_m, \dots, x_n, y_n - 1$ to \mathbb{R} , we have that $|\nabla \hat{\zeta}_m^n(x, y)|^2 \leq c' \log n$ with c' depending on δ . Lemma 11 then implies that

$$P\left(|\hat{\zeta}_m^n - E\hat{\zeta}_m^n| > n^{-1/4}\right) \leq 2e^{-c''(n^{1/2}/\log n)}, \quad (4.1)$$

for all $m > \delta n$ with a $c'' = c''(V, \beta, a, \delta)$. But then by (3.13) of Corollary 15 have that $E\hat{\zeta}_m^n = O((\log n/n)^{1/2})$ uniformly in $m > \delta n$, and the same estimate will hold for $E\zeta_m^n$. Now the result follows from (4.1) and a union bound. \square

4.2 Finite dimensional convergence

We employ a classical blocking argument, with the limit being understood through the sum over “good blocks” (of length $O(n^{1/6})$) of the variables, each such block separated by $O(\log n)$ “buffers”. That the minimizer of any conditional Hamiltonian becomes independent of the boundary in $O(\log n)$ steps will produce the required decorrelation between adjacent blocks.

Define recursively the times,

$$m_1 = 1, \quad m_{k+1} = \begin{cases} m_k + \lfloor c \log n \rfloor & \text{if } k \text{ is odd} \\ m_k + \lfloor n^{1/6} \rfloor & \text{if } k \text{ is even,} \end{cases}$$

and corresponding good blocks and buffers: for $i = 1, 2, \dots$,

$$\begin{aligned} \mathcal{G}_i &= \frac{1}{\phi(m_{2i}/n)} \sum_{k=m_{2i}}^{m_{2i+1}} [(X_k - x_k^o) - (Y_k - y_k^o)], \\ \mathcal{B}_i &= \frac{1}{\phi(m_{2i-1}/n)} \sum_{k=m_{2i-1}}^{m_{2i}} [(X_k - x_k^o) - (Y_k - y_k^o)]. \end{aligned} \quad (4.2)$$

Here we made one more approximation in pulling the ϕ^{-1} of smallest index out of each block sum. By the continuity of ϕ it will be clear that this will make no difference in what follows. Also, truncating the final \mathcal{G}_i sum if necessary we can always assume that the stretch $[n - c \log n, n]$ is buffer. With this setup the result is:

Lemma 18. *Set $\mathcal{G}_t^n = \sum_{i:m_i \in [nt, n]} \mathcal{G}_i$ and $\mathcal{B}_t^n = \sum_{i:m_i \in [nt, n]} \mathcal{B}_i$. Then as $n \rightarrow \infty$, there is a suitable choice of $c = c(V, \beta, a)$ in (4.2) so that for any k and $0 < t_1 < t_2 < \dots < t_k \leq 1$, $(\mathcal{G}_{t_1}^n, \mathcal{G}_{t_2}^n, \dots, \mathcal{G}_{t_k}^n)$ and $(\mathcal{B}_{t_1}^n, \mathcal{B}_{t_2}^n, \dots, \mathcal{B}_{t_k}^n)$, converge in law to a centered Gaussian vector with covariance $\frac{1}{\beta} \log \frac{1}{\theta(t_i)} \wedge \frac{1}{\beta} \log \frac{1}{\theta(t_j)}$ and the zero vector, respectively.*

Proof. We start by estimating $Ee^{i\tau \mathcal{G}_t^n}$ for t fixed. With I_i the support of any corresponding \mathcal{G}_i figuring into \mathcal{G}_t^n , denote by K_i the interval formed by adjoining the $(c/3) \log n$ length

stretches of indices to the left/right of I_i . The parameter c is chosen large enough so that the strategy of Proposition 13 can be followed (the length $c \log n$ buffer about I_i is now length $(c/3) \log n$, but c is chosen as needed in both cases). In particular, boundary values set at ∂K_i will have weak influence on the statistics of \mathcal{G}_i .

With q_i the variables in ∂K_i we by now understand that

$$E e^{i\tau \mathcal{G}_i^n} = E \left[\mathbf{1}_Q \prod_{i: m_i \in [nt, n]} E_{q_i} [e^{i\tau \mathcal{G}_i}] \right] + o(1) \quad (4.3)$$

where Q is the event that $\max_i \|q - (x^o, y^o)\|_{\partial K, \infty} \leq c' \sqrt{\frac{\log n}{n}}$ for suitably large c' . Further,

$$\left| E_{q_i} [e^{i\tau \mathcal{G}_i} - (1 + i\tau \mathcal{G}_i - \frac{1}{2} \tau^2 \mathcal{G}_i^2)] \right| \leq E_{q_i} |\tau \mathcal{G}_i|^3 = O(|I_i|^3 n^{-3/2}),$$

by Proposition 13 (or rather its proof). Similarly, by Corollary 15 we have that $E_{q_i} \mathcal{G}_i = O(|I_i|(\log n)^5 n^{-3/2})$. Combining these facts with Corollary 14 we find that

$$E_{q_i} e^{i\tau \mathcal{G}_i} = 1 - \frac{\tau^2}{2} \frac{\theta'(m_{2i}/n)}{\beta \theta(m_{2i}/n)} \frac{(m_{2i+1} - m_{2i})}{n} + \kappa_n \quad (4.4)$$

with $|\kappa_n| = o((m_{2i+1} - m_{2i})n^{-1})$ uniformly in $q_i \in Q$. Substituting back into (4.3) we recognize the Riemann sum for $\int_t^1 \theta'(s) \theta^{-1}(s) ds$ on scale $\Delta = (m_{2i+1} - m_{2i})n^{-1}$.

The same considerations apply to \mathcal{B}_t^n , with the right hand side of (4.4) modified by shifting $2i$ to $2i - 1$. But $(m_{2i} - m_{2i-1})n^{-1} = o(\Delta)$ while there are still Δ^{-1} factors in the analog of (4.3). The outcome is that $E e^{i\tau \mathcal{B}_t^n} \rightarrow 1$ as $n \rightarrow \infty$. To be precise we note that while Corollaries 14 and 15 do not apply to the final block $\mathcal{B}_{i(n)}$ in \mathcal{B}_t^n (as it is constructed to be supported on $[n - c \log n, n]$), a more crude estimate by Proposition 13 gives that $E(\mathcal{B}_{i(n)})^2 = o(1)$.

The convergence of k -point marginals follows from the asymptotic independence of increments for $t \mapsto \mathcal{G}_t^n$ which is immediate from the necessary version of (4.3). Taking $k = 2$ gets the point across. With $s < t$ and any τ and ν ,

$$\begin{aligned} E e^{i\tau \mathcal{G}_s^n + i\nu \mathcal{G}_t^n} &= E [E_q e^{i\tau(\mathcal{G}_s^n - \mathcal{G}_t^n)} E_{q'} e^{i(\tau + \nu)\mathcal{G}_t^n}] \\ &= e^{-\frac{\tau^2}{2\beta} \log \frac{\theta(t)}{\theta(s)}} e^{-\frac{(\tau + \nu)^2}{2\beta} \log \frac{1}{\theta(t)}} + o(1), \end{aligned}$$

and the exponent reads $(-\frac{1}{2\beta}) \times (\tau^2 \log \frac{1}{\theta(s)} + \nu^2 \log \frac{1}{\theta(t)} + 2\tau\nu \log \frac{1}{\theta(t)})$ as desired. In line one above we simply use that no \mathcal{G}_i is included in both $(\mathcal{G}_s^n - \mathcal{G}_t^n)$ and \mathcal{G}_t^n – the convention being it belongs to the sum in which its left-most point of support lies. Thus the conditionings variables q and q' can be chosen not to overlap, and to be a distance $O(\log n)$ from any of the corresponding \mathcal{G}_i 's within. Now we simply apply the strategy inherent in (4.4) to the E_q and $E_{q'}$ expectations separately. \square

4.3 Pathwise convergence

To lift the convergence from marginal distributions to convergence in the space of continuous paths we show the following.

Lemma 19. *With ζ_t^n either equal to \mathcal{G}_t^n or \mathcal{B}_t^n defined in the statement of Lemma 18 and all $0 \leq r \leq s \leq t \leq 1$:*

$$E [(\zeta_r^n - \zeta_s^n)^2 (\zeta_s^n - \zeta_t^n)^2] \leq c \left(\log \frac{\theta(t)}{\theta(r)} \right)^2 \quad (4.5)$$

for a constant c and all large enough n .

This suffices for the tightness of $t \mapsto \mathcal{G}_t^n$ and $t \mapsto \mathcal{B}_t^n$ due to Theorem 13.5 of [4]. To compare the above with that statement, note that we are using the latter in the case that the limit process is continuous, the α and β parameters defined there equal to one, and choice of $F(t) = -\sqrt{c} \log \theta(t)$ (our time-like parameter naturally runs “in reverse”). We can then conclude the full convergence of the desired process $t \mapsto \mathcal{G}_t^n + \mathcal{B}_t^n$ to $t \mapsto \int_{\theta(t)}^1 \frac{db_u}{\sqrt{\beta u}}$ by Slutsky’s Lemma.

Proof. We show the inequality (4.5) holds for the process of good blocks \mathcal{G}_t^n . A similar calculation will apply to \mathcal{B}_t^n . To start write

$$E [(\mathcal{G}_r^n - \mathcal{G}_s^n)^2 (\mathcal{G}_s^n - \mathcal{G}_t^n)^2] = \sum_{\substack{i_1, i_2: m_{2i} \in [nr, ns] \\ j_1, j_2: m_{2j} \in [ns, nt]}} E[\mathcal{G}_{i_1} \mathcal{G}_{i_2} \mathcal{G}_{j_1} \mathcal{G}_{j_2}], \quad (4.6)$$

recalling (4.2). The main contribution stems from terms on the right hand side of (4.6) in which $i_1 = i_2$ and $j_1 = j_2$. For any such term we have that: with $\delta_n = c' \sqrt{\frac{\log n}{n}}$ and large enough c' ,

$$\begin{aligned} E[\mathcal{G}_i^2 \mathcal{G}_j^2] &= E \left[E_{q_i}[\mathcal{G}_i^2] E_{q_j}[\mathcal{G}_j^2], \|q_i, q_j\|_\infty \leq \delta_n \right] + o(n^{-2}) \\ &\leq c'' \frac{(m_{2i+1} - m_{2i})(m_{2j+1} - m_{2j})}{n^2 \phi(m_{2i}/n) \phi(m_{2j}/n)} + o(n^{-2}), \end{aligned} \quad (4.7)$$

by reasoning used several times before. And as in all such cases, we can choose the supports of the disjoint boundary q_i and q_j a large multiple of $\log n$ away from the respective supports of \mathcal{G}_i and \mathcal{G}_j . The inequality in (4.7) is then a direct consequence of Corollary 14. Summed over all $O(n^2 |m_{2k+1} - m_{2k}|^{-2}) = O(n^{5/3})$ possible i and j we get a constant multiple of $\int_r^s \frac{du}{\phi(u)} \int_s^t \frac{dv}{\phi(v)}$ upper bound for the corresponding subsum of (4.6).

Terms of type $i_1 \neq i_2$ or $j_1 \neq j_2$ in (4.6) are easily seen to be subdominant given, in this regime of indices, $|E_q[\mathcal{G}_i]| \mathbf{1}_{\|q\|_\infty \leq \delta_n} = O(n^{-(3/2-)}).$ This is a byproduct of the proof of Corollary 15. \square

5 Convergence in norm

The results of the previous section imply the pointwise convergence of $K_n(s, t)$ to $K(s, t)$, at least over subsequences on a suitable probability space. The proof of the convergence of the corresponding operators in Hilbert-Schmidt norm (and in the same subsequential coupling) would follow if we could build a dominating kernel \widehat{K} (that is, $K_n(s, t) \leq \widehat{K}(s, t)$) which lies almost surely in $L^2([0, 1]^2)$. Note that one readily checks that $\int_0^1 \int_0^t |K(s, t)|^2 ds dt < \infty$ with probability one.

The next proposition provides such an estimate, but only away from the singularity at the origin.

Proposition 20. *For sufficiently large $c = c(V, \beta, a)$ and any $\epsilon > 0$,*

$$K_n(s, t) \leq C_n \frac{(\theta(s)\theta(t))^{-\epsilon}}{(\phi(s)\phi(t))^{1/2}} \left(\frac{\theta(s)}{\theta(t)} \right)^{\frac{a}{2} + \frac{1}{4}} \quad \text{for } c \frac{\log n}{n} \leq s \leq t \leq 1, \quad (5.1)$$

in which $C_n = C_n(c, \epsilon)$ is a tight random sequence.

The point is that, denoting the deterministic part of the right hand side of (5.1) by

$$K_\epsilon(s, t) = \frac{(\theta(s)\theta(t))^{-\epsilon}}{(\phi(s)\phi(t))^{1/2}} \left(\frac{\theta(s)}{\theta(t)} \right)^{\frac{a}{2} + \frac{1}{4}}, \quad (5.2)$$

a calculation shows that $\int_0^1 \int_0^t |K_\epsilon(s, t)|^2 ds dt < \infty$ as long as $\epsilon < \frac{1}{4} \wedge \frac{(a+1)}{2}$. To bridge the gap for small values of s and t , we will show the following.

Proposition 21. *For any $c > 0$,*

$$\iint_{0 \leq s \leq t \leq 1, s \leq c \frac{\log n}{n}} |K_n(s, t)|^2 ds dt \rightarrow 0 \quad (5.3)$$

in probability.

We mention that in establishing the convergence of the classical β -Laguerre matrix model to $\text{SBO}_{\beta, a}$ in [18] a single dominating kernel was relatively easy to come by. On the other hand, for the “spiked” hard edge considered in [19] in which case one deals with matrix kernel operators a similar cutoff procedure was required $O(\log n)$ steps away from the singularity.

In any case, one may now argue as follows. Given any subsequence of operators K_n , choose a further subsequence $K_{n'}$ and a probability space on which (5.3) takes place almost surely and the bound (5.1) holds almost surely with the tight sequence $C_{n'}$ replaced by some deterministic constant (bounding the chosen subsequential limit of tight the prefactors). Presuming the pointwise convergence $K_{n'} \rightarrow K$ of kernels also takes place almost surely

on the same space (which may be achieved by taking yet a further subsequence), it follows that $\int_0^1 \int_0^t |K_{n'}(s, t) - K(s, t)|^2 ds dt \rightarrow 0$ with probability one. This completes the proof of Theorem 2 and hence the main result.

The proofs of Propositions 20 and 21 occupy the next two subsections.

5.1 Tight kernel bound away from the singularity

Proposition 20 is a consequence of the following.

Lemma 22. *Define $h(t) = (1 + \log \frac{1}{\theta(t)})^p$. Then for any $p \in (1/2, 1)$ and $c = c(V, \beta, a)$ large enough, the sequence*

$$\max_{k \geq c \log n} \frac{1}{h(k/n)} \sum_{j=k}^n \log \left(\frac{Y_j/y_j^o}{X_j/x_j^o} \right) \quad (5.4)$$

is tight.

The results of Sections 2 and 3 (concentration about the minimizers and proximity of the coarse and true minimizers) show that $\max_{k > c \log n} (\phi(k/n)/X_k)$ is also tight, controlling the prefactor to the random product appearing in the definition of $K_n(s, t)$. In particular then there is a tight random sequence C_n such that $1/X_k \leq C_n/\phi(k/n)$ and $\sum_{j=k}^n \log(\frac{Y_j/y_j^o}{X_j/x_j^o}) \leq C_n h(k/n)$, at least for k in the prescribed range. But that means that

$$\begin{aligned} K_n(s, t) &= X_{[nt]}^{-1} e^{\sum_{k=[ns]}^{[nt]-1} \log \left((Y_k/y_k^o)/(X_k/x_k^o) \right)} e^{-\sum_{k=[ns]}^{[nt]-1} \log \left(x_k^o/y_k^o \right)} \\ &\leq C_n \frac{e^{C_n(h(s)+h(t))}}{(\phi(s)\phi(t))^{1/2+\epsilon}} \left(\frac{\theta(s)}{\theta(t)} \right)^{\frac{a}{2}+\frac{1}{4}}, \quad \text{for } s, t > c \frac{\log n}{n}. \end{aligned} \quad (5.5)$$

Here we have used that:

$$\sum_{j=[nt]}^n \log(x_j^o/y_j^o) \leq (a/2 + 1/4) \log \theta(t) - 1/2 \log \phi(t) + c'(1 + \sqrt{\log \phi(t)})$$

for $nt > c \log n$ and some constant c' , recall Proposition 8. This last (error) term is then absorbed into the $C_n \phi(t)^\epsilon$ in (5.5). That estimate in turn implies the claimed inequality (5.1) of Proposition 20 as for any positive c and ϵ and $p \in [0, 1)$ there is a $c'' = c(\epsilon, p)$ so that $c(1+a)^p \leq c' + \epsilon a$ for all $a \geq 0$. Afterwards ϵ is adjusted. Note here and below we write $K_n(s, t) = X_{[nt]}^{-1} \prod_{k=[ns]}^{[nt]-1} (Y_k/X_k)$. This is not quite accurate, compare the definition (1.11), but can be considered a convenient shorthand and will not make any difference for the level of estimate required here.

As for Lemma 22 we need one last ingredient. This is due to Dudley [8] (though see Proposition 2.2.10 of [20] for a succinct proof).

Proposition 23. Consider a metric space (T, d) and a centered process $(Z_t)_{t \in T}$ with law \mathbf{P} satisfying

$$\mathbf{P}(|Z_s - Z_t| \geq \lambda) \leq 2e^{-\frac{\lambda^2}{2d(s,t)^2}} \quad (5.6)$$

for all $\lambda > 0$. Then there is a universal constant c such that

$$\mathbf{E} \sup_{t \in T} Z_t \leq c \sum_{q \geq 0} 2^{q/2} e_q(T), \quad (5.7)$$

in which $e_q(T) = \inf \sup_{t \in T} d(t, T_q)$ and the infimum is over all $T_q \subset T$ of cardinality $\leq 2^{2^q}$.

Proof of Lemma 22. The first step is to truncate the logarithm. With $c' > 1$ to be chosen momentarily let $\delta_n = c' \sqrt{\frac{\log n}{n}}$, and for each $j \in [c \log n, n]$, where $c \geq c'$ will also be chosen along the way, define:

$$G_j(z) = \begin{cases} \log(z/z_j^o), & \text{for } |z - z_j^o| \leq \delta_n, \\ \log(1 - \delta_n/z_j^o) \text{ or } \log(1 + \delta_n/z_j^o), & \text{for } z \leq z_j^o - \delta_n \text{ or } z \geq z_j^o + \delta_n. \end{cases} \quad (5.8)$$

Here z_j^o denotes the coordinate of the minimizer x_j^o or y_j^o according to whether G_j is to be evaluated at x_j or y_j . Since both x_j^o and y_j^o can be bounded below by a small constant multiple of $\phi(j/n)$ which in turn is $O(\sqrt{j/n})$ for small j , by choice of $c = c(c')$ we have that:

$$\delta_n/x_j^o \vee \delta_n/y_j^o \leq \frac{1}{2}, \quad (x_j^o - \delta_n) \wedge (y_j^o - \delta_n) \geq \frac{1}{2} \sqrt{\frac{j}{n}}, \text{ for all } j \geq c \log n. \quad (5.9)$$

These bounds at least guarantee that (5.8) is sensible. Further, for all k in the range of interest the sum

$$S_k = S_k(X, Y) = \sum_{j=k}^n G_j(Y_j) - G_j(X_j) \quad (5.10)$$

agrees with $\sum_{j=k}^n \log \left(\frac{Y_j/y_j^o}{X_j/x_j^o} \right)$ on the event $Q = \{|X_j - x_j^o|, |Y_j - y_j^o| \leq \delta_n \text{ for all } j \geq \log n\}$. And then once again Proposition 9 along with a union bound implies $P(Q) = 1 - o(1)$ granted $c' = c'(V, \beta, a)$ is chosen large enough. Hence, to prove the claim it suffices to show that

$$\max_{k \geq c \log n} \frac{S_k}{h(k/n)} \text{ is tight.} \quad (5.11)$$

Where now $c = c(V, \beta, a)$ is fixed after an appropriate choice of c' .

Next we note that, as a map taking $(x_k, \dots, x_n; y_k, \dots, y_n) \mapsto \mathbb{R}$, S_k has square Lipschitz norm bounded as in

$$\|\nabla S_k(x, y)\|_2^2 \leq \sum_{j=k}^n \frac{1}{(x_j^o - \delta_n)^2} + \frac{1}{(y_j^o - \delta_n)^2} \leq 8n \sum_{j=k}^n \frac{1}{j} \leq 16n \log(n/k), \quad (5.12)$$

where the second inequality in (5.9) is used. Therefore, by Lemma 11

$$P\left(S_k(X, Y) \geq ES_k(X, Y) + \lambda\right) \leq \exp\left(-\frac{\lambda^2}{c'' \log \frac{n}{k}}\right), \quad (5.13)$$

for all $\lambda > 0$ and all $k \geq c \log n$ with yet another constant $c'' = c''(V, \beta, a)$

Turning to (5.11), we introduce

$$F_m = \max_{e^{-m-1} < \frac{k}{n} \leq e^{-m}} S_k, \quad \text{for } m = 1, 2, \dots, \quad (5.14)$$

and, noting that it is only the small (as in $o(n)$) values of k which really require attention, estimate as follows:

$$\begin{aligned} P\left(\max_{c \log n \leq k \leq n/4} \frac{S_k}{h(k/n)} > \lambda\right) &\leq \sum_{m=1}^{\infty} P(F_m > \lambda h(e^{-m})) \\ &\leq \sum_{m=1}^{\infty} \exp\left(-(\lambda m^p - EF_m)^2 / 4c''m\right). \end{aligned} \quad (5.15)$$

The second inequality is due to (5.13) along with the fact that the maximum function has Lipschitz norm one. Recalling that $p > 1/2$, the proof will be finished by showing that EF_m is uniformly bounded in m and n (in which case the right hand side above can be made as small as one likes by taking $\lambda \uparrow \infty$). This is where Proposition 23 comes in.

Another application for (5.13) shows that the condition (5.6) of that Proposition is satisfied with the discrete process $k \mapsto S_k - ES_k$ in the role of $t \mapsto Z_t$ with $T = [ne^{-m-1}, ne^{-m}]$ and metric $d(k, \ell)$ equal to a constant (*i.e.*, independent of both n and m) multiple of $\sqrt{|\log k/\ell|}$. In particular, $T = T(m, n)$ has diameter bounded independently of n or m . Thus, the $e_q(T)$ in the punchline (5.7) can be bounded by $\sup_{k \in T} d(k, T_q)$ for an equally spaced T_q , with the result that $e_q = O(2^{-2^q})$ and

$$E \left[\max_{e^{-m-1} \leq k/n \leq e^{-m}} (S_k - ES_k) \right] \leq c''',$$

where c''' can be chosen fixed for all n and m .

It is left is to demonstrate that $\sup_{k > c \log n} ES_k$ is similarly bounded. Since $E \log(Z_k/z_k^o) \mathbf{1}_{Q^c}$ (for (Z, z) either (X, x) or (Y, y)) can made exponentially small in n , we have

$$E \log \left(\frac{Y_k/y_k^o}{X_k/x_k^o} \right) = E \left(\frac{X_k - x_k^o}{x_k^o} - \frac{Y_k - y_k^o}{y_k^o} \right) + E \left(\frac{(X_k - x_k^o)^2}{2(x_k^o)^2} - \frac{(Y_k - y_k^o)^2}{2(y_k^o)^2} \right) + O(k^{-3/2}),$$

after restoring the integrals to the full domain (from Q) and using that $E|Z_k - z_k^o|^3 = O(n^{-3/2})$ and once more that $z_k^o \geq c\sqrt{k/n}$ in the last term. But by Corollary 15 we have for example,

$$\begin{aligned} \sum_{k=c \log n}^n \left| E \left[\frac{X_k - x_k^o}{x_k^o} - \frac{Y_k - y_k^o}{y_k^o} \right] \right| &= \sum_{k=c \log n}^n \phi(k/n)^{-1} |E(X_k - x_k^o) - (Y_k - y_k^o)| (1 + o(1)) \\ &\leq c'' \sum_{k=c \log n}^n (\log k)^5 k^{-3/2} = o(1), \end{aligned}$$

with a similar conclusion for the sum of the mean-square differences. \square

5.2 Near the singularity

Proposition 21 actually uses Proposition 20 as input, in addition to the next rough estimate. Again the strategy is similar to that in [19] (see Sections 3.5-3.6 there).

Lemma 24. *For any $c > 0$ there exist events \mathcal{B}_n of probability tending to one on which*

$$K_n(s, t) \leq \frac{c'}{\phi(t)} \exp \left(\kappa_n \int_s^t \frac{d\tau}{\phi(\tau)} \right) \quad \text{for } 0 \leq s \leq t \leq c \frac{\log n}{n}, \quad (5.16)$$

with a constant $c' = c(V, \beta, a, c)$ and $\kappa_n = c' \sqrt{n \log_2 n}$.

Granted this we will first prove the proposition and return to the proof of Lemma 24 afterwards.

Proof of Proposition 21. As the integral in question in (5.3) is increasing in c we may as well assume that it is large enough so that Proposition 20 is in place. In addition, we will invoke that proposition in the following way. Denote by \mathcal{A}_n the event that the random sequence C_n appearing (5.1) exceeds some c' . By choice of c' we can take the probability of \mathcal{A}_n as close to one as we like. We are left to show that the appraisal (5.3) takes place on the intersection of \mathcal{A}_n and \mathcal{B}_n .

Now denoting $\delta_n = c \frac{\log n}{n}$ and on \mathcal{B}_n , Lemma 24 gives that,

$$\begin{aligned} \iint_{0 \leq s \leq t \leq \delta_n} |K_n(s, t)|^2 ds dt &\leq c' \int_0^{\delta_n} \int_0^t \frac{1}{t} e^{\kappa_n(\sqrt{t} - \sqrt{s})} ds dt \\ &= 2c' \int_0^{\delta_n} \left(\frac{e^{\kappa_n \sqrt{t}} - 1}{\kappa_n^2 t} - \frac{1}{\kappa_n \sqrt{t}} \right) dt \\ &\leq c' \frac{e^{\kappa_n \sqrt{\delta_n}}}{\kappa_n^2}, \end{aligned} \quad (5.17)$$

which is $O(\frac{1}{n^{1-\eta}})$ for any $\eta > 0$. The first inequality uses (5.16) along with the fact that $\phi(t)$ is bounded above and below by constant multiples of \sqrt{t} for $t \ll 1$ (and absorbs these constants into an adjusted c'). The second inequality is an elementary Laplace approximation.

On the remaining domain of integration write

$$\int_{\delta_n}^1 \int_0^{\delta_n} |K_n(s, t)|^2 ds dt = \int_{\delta_n}^1 |K_n(\delta_n, t)|^2 dt \int_0^{\delta_n} X_{[n\delta_n]}^2 |K_n(s, \delta_n)|^2 ds,$$

which, restricted to the event $\mathcal{A}_n \cap \mathcal{B}_n$, is bounded above as in

$$\begin{aligned} \int_{\delta_n}^1 \int_0^{\delta_n} |K_n(s, t)|^2 ds dt &\leq c' \int_{\delta_n}^1 |K_n(\delta_n, t)|^2 dt \int_0^{\delta_n} e^{\kappa_n(\sqrt{\delta_n} - \sqrt{s})} ds \\ &\leq c' \left(\delta_n^{a-2\epsilon} \int_{\delta_n}^1 t^{-(a+1+2\epsilon)} dt \right) \frac{e^{\kappa_n \sqrt{\delta_n}}}{\kappa_n^2}. \end{aligned} \quad (5.18)$$

The first line employs the same arguments used in the first line of (5.17), as well as the observation that the proof of Lemma 24 includes the bound $X_{[n\delta_n]}K_n(s, \delta_n) \leq c'e^{\kappa_n \int_s^t \frac{d\tau}{\phi(\tau)}}$ on the event in question. The second line recalls the definition of K_ϵ from (5.2) and again that uses the largest contribution comes from of the origin where $\theta(t)$ and $\phi^2(t)$ are bounded in terms of t . Bounding the remaining integral gives that (5.18) is controlled by a constant multiple of $\delta_n^{-4\epsilon} \kappa_n^{-2} e^{\kappa_n \sqrt{\delta_n}}$, which tends to zero like a small negative power of n by choosing $\epsilon > 0$ small enough. \square

It remains to go back and establish Lemma 24.

Proof of Lemma 24. The events \mathcal{B}_n are constructed so that the inequality

$$\frac{Y_k}{X_k} \leq 1 + c' \frac{\sqrt{\log_2 n}}{\sqrt{n}\phi(k/n)} \quad (5.19)$$

holds for all indices $k \leq c \log n$ with a fixed constant $c' = c'(c, \beta, a, V)$. Granted this one has that

$$X_{[nt]}K_n(s, t) = \prod_{k=[ns]}^{[nt]-1} \frac{Y_k}{X_k} \leq \exp \left(c' \sqrt{n \log_2 n} \sum_{k=[ns]}^{[nt]-1} \frac{1}{n\phi(k/n)} \right),$$

for $s, t \leq c \frac{\log n}{n}$, simply due to $(1+a) \leq e^a$ for $a \geq 0$. Further estimating above the obvious Riemann sum produces the exponential factor in the advertised (5.16). An appropriate upper bound on the $(X_{[nt]})^{-1}$ prefactors will follow in the course of establishing (5.19).

To begin, using Proposition 9 yet again we have that

$$P \left(|(X_k, Y_k) - (x_k^o, y_k^o)| \geq c' \sqrt{\frac{\log_2 n}{n}} \text{ for any } k \leq c \log n \right) \leq c \log n \times (\log n)^{-\gamma(c')}, \quad (5.20)$$

where γ can be made large by choice of c' . On the other hand we also have that

$$|x_k^o - \phi(k/n)| + |y_k^o - \phi(k/n)| \leq c' \frac{1}{\sqrt{n}}, \quad (5.21)$$

for $k \leq c \log n$. The latter follows from the established $O(1/\sqrt{n})$ closeness of the true and fine minimizers for $k \leq c \log n$ (Proposition 8) coupled with the explicit formulas for the fine minimizers. Now set

$$\mathcal{B}_n^{(1)} = \left\{ |(X_k, Y_k) - (\phi(k/n), \phi(k/n))| \leq c' \sqrt{\frac{\log_2 n}{n}} \text{ for } c'' \log_2 n \leq k \leq c \log n \right\}$$

with a c'' to be chosen momentarily. We have just explained why $P(\mathcal{B}_n^{(1)}) = 1 - o(1)$, while

on that event there is the bound

$$\begin{aligned} \frac{Y_k}{X_k} &\leq \frac{\phi(k/n) + c' \sqrt{\frac{\log_2 n}{n}}}{\phi(k/n) - c' \sqrt{\frac{\log_2 n}{n}}} \\ &\leq 1 + 4c' \frac{\sqrt{\log_2 n}}{\sqrt{n}\phi(k/n)}, \quad \text{granted that } c' \frac{\sqrt{\log_2 n}}{\sqrt{n}\phi(k/n)} < \frac{1}{2}, \end{aligned} \quad (5.22)$$

which can be guaranteed by taking c'' large depending on c' . This is (5.19), after a readjustment of c' .

Moving to the range $k \leq c'' \log_2 n$ first observe that for such indices the right hand side of (5.19) can be replaced by a constant multiple of $\sqrt{\frac{\log_2 n}{k}}$. Here we select a small $\delta > 0$ such that

$$\delta \sqrt{\frac{k}{n}} \leq \phi(k/n) \leq \frac{1}{\delta} \sqrt{\frac{k}{n}} \quad \text{for } k \leq c'' \log_2 n, \quad (5.23)$$

and define

$$\begin{aligned} \mathcal{B}_n^{(2)} &= \left\{ |Y_k - \phi(k/n)| \leq c' \sqrt{\frac{\log_3 n}{n}} \text{ for } k \leq c'' \log_2 n \right\} \\ &\cap \left\{ X_k > \delta^2 \phi(k/n) \text{ for } \log_4 n \leq k \leq c'' \log_2 n, X_k > \frac{1}{\sqrt{n \log_3 n}} \text{ for } k \leq \log_4 n \right\}. \end{aligned}$$

With the corresponding restrictions on X_k and Y_k in place it holds that

$$\frac{Y_k}{X_k} \leq \frac{1}{\delta^2} + \frac{c'}{\delta} \sqrt{\frac{\log_3 n}{k}}, \quad \text{for } k \in [\log_4 n, c'' \log_2 n], \quad (5.24)$$

while

$$\frac{Y_k}{X_k} \leq \left(\frac{1}{\delta} + c' \right) \log_3 n, \quad \text{for } k \in [1, \log_4 n]. \quad (5.25)$$

The right hand sides of both (5.24) and (5.25) are then $O\left(\sqrt{\frac{\log_2 n}{k}}\right)$ as desired.

Leaving aside the verification that $P(\mathcal{B}_n^{(2)}) = 1 - o(1)$, the claim is that the proof is complete by choosing $\mathcal{B}_n = \mathcal{B}_n^{(1)} \cap \mathcal{B}_n^{(2)}$. The remaining detail is the prefactor $[X_{\lfloor nt \rfloor}]^{-1}$ multiplying $\prod_{k=\lfloor ns \rfloor}^{\lfloor nt \rfloor - 1} \frac{Y_k}{X_k}$ in the definition of the kernel. For $k = \lfloor nt \rfloor \geq \log_4 n$ the definition of \mathcal{B}_n explicitly restricts $[X_k]^{-1}$ to be less than a constant multiple of $[\phi(k/n)]^{-1}$, as desired. For smaller values of k the bound available from the definition of $\mathcal{B}_n^{(2)}$ is off by a factor of $\sqrt{\frac{\log_3 n}{k}}$. But this is readily absorbed into the upper bound on Y_k/X_k provided by (5.25).

Returning to the probability of $\mathcal{B}_2^{(n)}$, that

$$P\left(|Y_k - \phi(k/n)| \geq c' \sqrt{\frac{\log_3 n}{n}} \text{ for } k \leq c'' \log_2 n\right) = o(1)$$

holds by the same reasoning behind (5.20) and (5.21). The twist is the different type of restriction placed on X_k from below for k in this range, the lower bound provided by Gaussian concentration now being cumbersome.

For $k \in [\log_4 n, c'' \log_2 n]$ we require an upper bound on

$$P(X_k \leq \delta^2 \phi(k/n)) \leq P(X_k \leq \delta \sqrt{k/n}).$$

As Lemma 25 below shows, this probability is less than a constant multiple of $P(\zeta \leq \epsilon E\zeta)$ in which $\zeta \sim \chi_{\beta(k+a)}$ and $\epsilon = \epsilon(\delta, a, \beta)$ can be taken less than $1/2$ granted that $\delta \ll 1$. Here we use that $k \gg 1$ and that for any χ_r random variable $\sqrt{r-1/2} \leq E\chi_r \leq \sqrt{r}$ as long as $r \geq 1$. Next bring in the following tail inequality: again with χ_r denoting a random variable of the indicated law and $r \geq 1$,

$$P(|\chi_r - E\chi_r| \geq \eta E\chi_r) \leq 2e^{-\eta^2 r/2}. \quad (5.26)$$

This is a consequence of the Logarithmic Sobolev Inequality for measures with strictly log-concave densities (Chapter 5 of [16]) along with the mentioned upper/lower bounds on $E\chi_r$. Combining these remarks the conclusion is that

$$P(X_k \leq \delta^2 \phi(k/n), \text{ for some } k \in [\log_4 n, c'' \log_2 n]) \leq c''' \sum_{k \geq \log_4 n} e^{-\frac{\beta}{8}k},$$

which tends to zero as $n \rightarrow \infty$.

Finally, for $k \leq \log_4 n$ (with the real problem being when k is order one) the inequality (5.26) becomes ineffective, but we get by with the more elementary $P(\chi_r \leq \delta) \leq \kappa \delta^r$ where κ is fixed (for $r \geq 1$). This simple estimate yields

$$P\left(X_k \leq \frac{1}{\sqrt{n \log_3 n}}, \text{ for some } k \leq \log_4 n\right) \leq \sum_{k=1}^{\log_4 n} \left(\frac{c'''}{\sqrt{\log_3 n}}\right)^{\beta(k+a)} \rightarrow 0,$$

after another application of Lemma 25. □

Lemma 25. *Let $k \leq c \log n$ and denote by I_k the interval $[(k-d) \vee 0, k+d]$. Then, for large enough n ,*

$$P(X_k \leq t \mid X_j, Y_j, j \in I_k) \mathbf{1}_{\{X_j, Y_j \leq c\sqrt{\frac{\log_2 n}{n}}, j \in I_k\}} \leq c' P(\zeta \leq c' \sqrt{nt}), \quad (5.27)$$

in which ζ is a $\chi_{\beta(k+a)}$ random variable and c' depends on c (and β, a, V).

Proof. First, conditioned on all the other variables X_k has density function proportional to,

$$f(x; z) = x^{\beta(k+a)-1} e^{-n\gamma x^2 + n\Gamma(x, z)},$$

where $\gamma > 0$ is a constant and $\Gamma(x, z)$ is polynomial in x and the other coordinates $z_j \in I_k$. Note that the exponent of any variable in Γ is at least two. Hence, with $p(t)$ denoting the left hand side of (5.27) we have that

$$p(t) \leq \frac{\int_0^t f(x; z) dx}{\int_0^{\frac{\log n}{\sqrt{n}}} f(x; z) dx} \mathbf{1}_{\{x, z \leq c\sqrt{\frac{\log_2 n}{n}}\}} \leq \frac{\int_0^t x^{\beta(k+a)-1} e^{-n\gamma(1-c''\frac{\log_2 n}{n})x^2} dx}{\int_0^{\frac{\log n}{\sqrt{n}}} x^{\beta(k+a)-1} e^{-n\gamma(1+c''\frac{\log_2 n}{n})x^2} dx},$$

with a constant c'' depending on V and β .

Next denote $c_+ = \gamma(1+c''\frac{\log_2 n}{n})$ and $c_- = \gamma(1-c''\frac{\log_2 n}{n})$. For large enough n there are the bounds $\gamma/2 \leq c_-$, $c_+ \leq 2\gamma$, while $(c_+/c_-)^r \leq 2$ for any exponent r of order $\log n$. Changing variables then produces

$$p(t) \leq 2 \frac{\int_0^{\sqrt{2\gamma n}t} x^{\beta(k+a)-1} e^{-x^2} dx}{\int_0^{\sqrt{\gamma/2} \log n} x^{\beta(k+a)-1} e^{-x^2} dx}.$$

The proof is completed by a standard stationary phase calculation which shows that the integral in the denominator can be bounded below (independently of n , $k \leq n$, β or a) by a constant multiple of $\int_0^\infty x^{\beta(k+a)-1} e^{-x^2} dx$, the regular $\chi_{\beta(k+a)}$ normalizer. \square

Appendix

We include here the derivation that our matrix model $B(X, Y)B(X, Y)^T$ (with (X, Y) sampled from the measure P) realizes the joint eigenvalue density (1.1). To simplify notation a bit we take $c \prod_{i < j} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^n \lambda_i^\gamma e^{-V(\lambda_i)}$ as the target density, with any $\gamma > -1$ and polynomial V .

Also, to make a more direct connection with the derivation of the β -Laguerre ensemble one finds in the literature (in say [9]) consider first an upper bidiagonal matrix M with coordinates labeled in decreasing order: $M_{i,i} = x_{n-i+1}$ for $i = 1, \dots, n$ and $M_{i,i+1} = y_{n-i}$ for $i = 1, \dots, n-1$, with all x_i, y_i positive. Also introduce the tridiagonal coordinates through a Jacobi matrix $T = T(a, b)$ with $T_{i,i} = a_{n-i+1}$ for $i = 1, \dots, n$ and $T_{i,i+1} = T_{i+1,i} = b_{n-i}$ for $i = 1, \dots, n-1$. Here each $a_i \in \mathbb{R}$ and each $b_i \in \mathbb{R}_+$. We track the calculation from eigenvalue/eigenvector coordinates to (x, y) coordinates via $Q\Lambda Q^\dagger = T = MM^T$. Here Q is the eigenvector matrix, of which we only need the first components. These can be chosen to be real positive, and are denoted (q_1, \dots, q_{n-1}) , noting that q_n is specified by $\sum_{i=1}^n q_i^2 = 1$.

Next, we have that the Jacobians for the maps from (λ, q) to (a, b) , and then from (a, b) to (x, y) are given by

$$J = q_n \frac{\prod_{i=1}^n q_i}{\prod_{i=1}^{n-1} b_i}, \quad J' = 2^n x_1 \prod_{i=2}^n x_i^2,$$

respectively. See [13, eq. 1.156] for the former. The latter is derived from the identities $a_i = x_i^2 + y_i^2$ and $b_i = x_{i+1}y_i$ (where $y_n = 0$ is understood). We will also need the well-known

relation,

$$\prod_{i < j} (\lambda_i - \lambda_j)^2 = \frac{\prod_{i=1}^{n-1} b_i^{2i}}{\prod_{i=1}^n q_i^2},$$

for which see [13, eq. 1.148].

Since we obviously have that $\sum_{i=1}^n V(\lambda_i) = \text{tr} V(MM^T)$, the necessary computation is:

$$\begin{aligned} \left(\prod_{i=1}^n \lambda_i \right)^\gamma \prod_{i < j} |\lambda_i - \lambda_j|^\beta (q_n^{-1} \prod_{i=1}^n q_i^{\beta-1}) dq \wedge d\lambda &= \left(\prod_{i=1}^n x_i^{2\gamma} \right) \left(\frac{\prod_{i=1}^{n-1} (x_{i+1} y_i)^{\beta i}}{\prod_{i=1}^{n-1} q_i^\beta} \right) J J' dx \wedge dy \\ &= 2^n \prod_{i=1}^n x_i^{2\gamma + \beta(i-1)+1} \prod_{i=1}^{n-1} y_i^{\beta i-1} dx \wedge dy. \end{aligned}$$

Putting in $\gamma = \frac{\beta}{2}(a+1) - 1$ we recognize the factors in $x_i^{\beta(a+i)-1}$ and $y_i^{\beta i-1}$ in the claimed bidiagonal matrix density (1.5). Here we have decided to work with $B = SMS^{-1}$ where S is the antidiagonal matrix of alternating signs. This transformation does not effect the joint density of the individual coordinates, and one has that the eigenvalues of BB^T and MM^T agree.

Acknowledgements. B.R. was supported in part by NSF grant DMS-1340489. It is a pleasure to thank Manjunath Krishnapur and Michel Ledoux for several helpful discussions.

References

- [1] Bekerman, F., Figalli, A., Guionnet, A., Transport maps for β -matrix models and universality, *Comm. Math. Phys.* **338** (2015), 589–619.
- [2] Bobkov, S.G., Ledoux, M., From Brunn-Minkowski to Brascamp-Lieb and to Logarithmic Sobolev inequalities. *GAFa* **10** (2000), 1028 – 1052.
- [3] Bourgade, P., Erdős, L., Yau, H.T., Edge universality of β -ensembles, *Comm. Math. Phys.* **332** (2014), 261–353.
- [4] Billingsley, P, *Convergence of Probability Measures*. Wiley Series in Probability and Statistics (2nd Edition), 1999.
- [5] Davis, C. All convex invariant functions of Hermitian matrices, *Arch. Math.* **8** (1957), 276–278.
- [6] Deift, P., Gioev, D. Universality at the edge of the spectrum for unitary, orthogonal and symplectic ensembles of random matrices. *Comm. Pure Appl. Math.* **60** (2007), no. 6, 867–910.

- [7] Deift, P. Gioev, D., Kriecherbauer, T., Vanlessen, M., Universality for orthogonal and symplectic Laguerre-type ensembles, *J. Stat. Phys.* **129** (2007), 949–1053.
- [8] Dudley, R.M., The sizes of compact subsets of Hilbert space and continuity of Gaussian processes, *J. Funct. Anal.* **1** 290–330 (1967), 290–330.
- [9] Dumitriu, I., Edelman, A., Matrix models for beta ensembles, *J. Math. Phys.* **43** (2002), 5830–5847.
- [10] Edelman, A., Sutton, B., From random matrices to stochastic operators, *J. Stat. Phys.* **127** (2007), 1121–1165.
- [11] Ercolani, N.M., McLaughlin, K.D. T-R., and Pierce, V.U. Random matrices, graphical enumeration, and the continuum limit of Toda lattices. *Comm. Math. Phys.* **278** (2008), 31–81.
- [12] Forrester, P.J. Exact results and universal asymptotics in the Laguerre random matrix ensemble. *J. Math. Phys.* **35** (1994), 2519–2551.
- [13] Forrester, P.J.. *Log-Gases and Random Matrices*. Princeton University Press, Princeton-Oxford, 2010.
- [14] Krishnapur, M., Rider B., Virág, B., Universality of the Stochastic Airy Operator, *Comm. Pure Appl. Math.* **69** (2016), 145–199.
- [15] Kuijlaars, A., Vanlessen, M., Universality for eigenvalue correlations at the origin of the spectrum, *Comm. Math. Phys.* **243** (2003), 163–191.
- [16] Ledoux, M. *The concentration of measure phenomenon*. Math. Surveys and Monograph **89** AMS (2001).
- [17] Ramírez, J., Rider, B., Virág, B., Beta ensembles, stochastic Airy spectrum and a diffusion, *J. Amer. Math. Soc.* **24** (2011), 919–944.
- [18] Ramírez, J., Rider, B. Diffusion at the random matrix hard edge, *Comm. Math. Phys.* **288** (2009), 887–906. (Erratum: *Comm. Math. Phys.* **307** (2011), 561–563.)
- [19] Ramírez, J., Rider, B. Spiking the random matrix hard edge, *To appear, PTRF (arXiv:1506:04988)*.
- [20] Talagrand, M. *Upper and lower bounds for stochastic processes*. Ergebnisse der Mathematik und ihrer Grenzgebiete. **60** Springer (2014).

- [21] Tracy, C., Widom, H. Level spacing distributions and the Bessel kernel,
Comm. Math. Phys. **161** (1994), 289–309.
- [22] Tracy, C., Widom, H. On orthogonal and symplectic matrix ensembles.
Comm. Math. Phys. **177** (1996) 727–754.